

VISIT...

LANZAROTE
Caliente.COM

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

Структура и биосинтез нуклеиновых кислот

*учебник
для вузов*



МОЛЕКУЛЯРНАЯ БИОЛОГИЯ

Структура и биосинтез нуклеиновых кислот

Под редакцией
академика А.С. Спирина

Допущено
Государственным комитетом СССР
по народному образованию
в качестве учебника для студентов
высших учебных заведений,
обучающихся по специальности
«Биология»



Москва
«Высшая школа» 1990

ББК 28.071
М 75
УДК 577.217

**В. И. Агол, А. А. Богданов, В. А. Гвоздев, А. И. Грагеров,
А. М. Колчинский, А. Д. Мирзабеков, В. Г. Никифоров**

Рецензенты:

филиал кафедры биохимии Киевского государственного университета имени Т. Г. Шевченко при Институте молекулярной биологии и генетики АН УССР (зав. филиалом кафедры — акад. АН УССР, проф. *Г. Х. Мацука*) и д-р биол. наук, проф. *Р. С. Шакулов* (зав. лабораторией молекулярной биологии ВНИИ Генетика)

Молекулярная биология: Структура и биосинтез нуклеиновых кислот: Учеб. для биол. спец. вузов/В. И. Агол, А. А. Богданов, В. А. Гвоздев и др.; Под ред. А. С. Спирина.— М.: Высш. шк., 1990.— 352 с.: ил.
ISBN 5-06-000086-9

В учебнике рассмотрены принципы структурной организации ДНК и РНК, механизмы репликации, репарации и рекомбинации ДНК, описаны механизмы транскрипции, посттранскрипционной модификации РНК и сплайсинга, особенности репликации и транскрипции геномов ДНК- и РНК-содержащих вирусов.

М 1903010000(4309000000)—064
001(01)—90 138—89

ББК 28.071
57А

Учебное издание

**Агол Вадим Израилевич, Богданов Алексей Алексеевич,
Гвоздев Владимир Алексеевич, Грагеров Александр Исаакович,
Колчинский Александр Маркович, Мирзабеков Андрей Дариевич,
Никифоров Вадим Георгиевич**

МОЛЕКУЛЯРНАЯ БИОЛОГИЯ:

Структура и биосинтез нуклеиновых кислот

Редактор М. М. Пенкина. Младшие редакторы Е. В. Бурова, Е. И. Попова.
Художник В. Н. Хомяков. Художественный редактор Т. А. Коленкова.
Технический редактор Н. А. Битюкова. Корректор Г. И. Кострикова.

ИБ № 7784

Изд. № Е-578. Сдано в набор 14.06.89. Подп. в печать 13.11.89. Т—17648.
Формат 60×88^{1/16}. Бум. офс. № 2. Гарнитура литературная. Печать офсетная. Объем 21,56 усл. печ. л. 21,56 усл. кр.-отт. 22,76 уч.-изд. л. Тираж 9000 экз. Заказ № 102.
Цена 1 р. 10 к.

Издательство «Высшая школа», 101430. Москва. ГСП-4, Неглинная ул., д. 29/14

Набрано в Ордена Октябрьской Революции и ордена Трудового Красного Знамени МПО «Первая Образцовая типография». Госкомпечати СССР 113054, Москва, Валовая, 28.
Отпечатано в Московской типографии № 8 Госкомпечати СССР. 101898, Москва, Центр, Хохловский пер., 7.

ISBN 5-06-000086-9

© Коллектив авторов, 1990

ПРЕДИСЛОВИЕ

Эта книга — первый том учебника по молекулярной биологии, который планируется издать в трех томах. Второй том этого учебника «Молекулярная биология: Структура рибосом и биосинтез белка», написанный А. С. Спириным, вышел в издательстве «Высшая школа» в 1986 г. Выход в свет третьего тома, посвященного структуре и функциям белков, планируется несколько позже.

Значительная часть настоящей книги написана по материалам курсов лекций по молекулярной биологии и молекулярной вирусологии, которые авторы на протяжении многих лет читают в МГУ им. М. В. Ломоносова. Одним из создателей лекционного курса молекулярной биологии в Московском университете был выдающийся советский биохимик и генетик Р. Б. Хесин, которому и посвящается эта книга.

Молекулярная биология — одна из наиболее бурно развивающихся дисциплин в современном естествознании. С каждым годом привносится множество новых фактов в эту область биологии, и почти каждый год совершаются открытия. Объем настоящего учебника слишком мал, чтобы вместить в себя огромный накопленный материал по структуре и биосинтезу нуклеиновых кислот. Поэтому в каждом разделе книги даны обобщенные и устоявшиеся представления о той или иной проблеме, проиллюстрированные примерами из современной литературы. Авторы не могли не воспользоваться также и тем обстоятельством, что в последнее пятилетие в нашей стране было издано несколько ценных (в основном переводных) учебных пособий по молекулярной биологии и смежным дис-

циплинам. Это позволило изложить здесь целый ряд вопросов в весьма кратком виде и порекомендовать читателю найти более полные сведения о них во вполне доступных изданиях.

В написании книги приняли участие: А. А. Богданов (Введение, гл. I); А. И. Грагеров (гл. II—VI); В. Г. Никифоров (гл. VII); В. А. Гвоздев (гл. VIII—XI); А. Д. Мирзабеков и А. М. Колчинский (гл. XII); В. И. Агол (гл. XIII).

Авторы с признательностью примут от читателей критические замечания и советы по совершенствованию данного учебника; их следует направлять в издательство «Высшая школа» по адресу: 101430, Москва, ГСП-4, Неглинная ул., 29/14.

Авторы

ВВЕДЕНИЕ

Молекулярная биология как самостоятельная наука, изучающая молекулярные основы жизнедеятельности клетки, возникла на рубеже 1940—1950 гг., когда была установлена генетическая роль дезоксирибонуклеиновых кислот (ДНК), а расшифровка структуры ДНК позволила описать в простых физико-химических терминах принцип передачи наследуемых признаков от родительской клетки к дочерним.

К этому времени история изучения нуклеиновых кислот насчитывала уже около восьмидесяти лет. Честь их открытия принадлежит выдающемуся швейцарскому биохимику Фридриху Мишеру, который в 1868—1872 гг. выделил из ядер клеток гноя и спермы лососа новое фосфорсодержащее вещество, названное им нуклеином (от греч. *nucleus* — ядро). Впервые нуклеиновую кислоту, свободную от белков, получил Р. Альтман в 1889 г., который и ввел этот термин в биохимию. Разработка методов выделения и изучение химического состава нуклеиновых кислот были продолжены в лабораториях А. Косселя, У. Джонса, П. Левина, О. Гаммерстена, Дж. Гулланда и др.

Усилиями этих ученых и их сотрудников удалось установить, что в природе существует два типа нуклеиновых кислот. Один из них содержит два пурина — аденин и гуанин, два пиримидина — цитозин и тимин, остатки дезоксипентозы и фосфорной кислоты. Другой вместо тимина содержит урацил, а вместо дезоксипентозы — пентозу. Так как дезоксипентозонуклеиновые кислоты (в современной терминологии — дезоксирибонуклеиновые кислоты, ДНК) выделяли в основном из тимуса теленка, а пентозонуклеиновые кислоты (рибонуклеиновые кислоты, РНК) — из дрожжей и растений, то долгое время существовала уверенность в том, что ядра клеток животных содержат только ДНК, а ядра клеток растений — только РНК. И лишь к середине 1930-х годов было доказано, что ДНК и РНК содержатся в каждой живой клетке. Первостепенная роль в утверждении этого фундаментального положения принадлежит А. Н. Белозерскому, впервые выделившему ДНК

из растений. С развитием методов цитохимии и гистохимии (Т. Касперссон, Ж. Браше и др.), а также методов фракционирования субклеточных структур (А. Даунс, А. Мирский, Дж. Паладе и др.) к концу 1940-х годов было установлено, что ДНК локализуется преимущественно в ядре, а РНК — в цитоплазме клеток.

К началу 1950-х годов были завершены работы по изучению принципов химического строения нуклеиновых кислот (А. Тодд, В. Кон и сотр.), когда было установлено строение их мономеров — нуклеозидов и нуклеотидов, и доказано, что и в ДНК, и в РНК нуклеотидные остатки связаны только 3'—5'-фосфодиэфирной связью. К этому же времени с помощью бумажной хроматографии были выяснены основные закономерности нуклеотидного состава ДНК и РНК (Э. Чаргафф и сотр.). В частности, было показано, что в ДНК аденин и тимин, гуанин и цитозин всегда содержатся в равных количествах; это имело принципиальное значение при установлении ее макромолекулярной структуры.

Выдающейся вехой в изучении нуклеиновых кислот стало открытие О. Эйвери и сотр. (1944). Они показали, что с помощью чистой ДНК наследуемый признак может быть перенесен из одной клетки в другую и что ДНК является носителем генетической информации. Это положение получило вскоре убедительное подтверждение в экспериментах А. Херши и М. Чейз с ДНК бактериофагов.

В 1953 г. Дж. Уотсон и Ф. Крик сумели правильно интерпретировать данные рентгеноструктурного анализа ДНК, накопленные в лабораториях Р. Франклин и М. Уилкинса, и на их основе построить модель пространственной структуры ДНК. Они показали, что макромолекула ДНК — это регулярная двойная спираль, в которой две полинуклеотидные цепи строго комплементарны друг другу. Из анализа модели следовало, что после расплетания двойной спирали на каждой из полинуклеотидных цепей может быть построена комплементарная ей новая, в результате чего образуются две дочерние молекулы, не отличимые от материнской ДНК. Через пять лет М. Мезельсон и Ф. Сталь экспериментально подтвердили этот механизм, а несколько раньше (1956) А. Корнберг открыл фермент ДНК-полимеразу, который на расплетенных цепях, как на матрицах, синтезирует новые, комплементарные им цепи ДНК.

Открытие генетической роли ДНК потребовало решения другой фундаментальной задачи — проблемы кода, с помощью которого нуклеотидный текст переводится на язык аминокислот — структурных единиц белка. Впервые эту задачу правильно сформулировал в начале 1950-х годов Г. Гамов, который предсказал, что этот код должен быть трехбуквенным и неперекрывающимся. Экспериментально общие свойства генетического кода были установлены Ф. Криком, С. Бреннером и сотр. к концу 1950-х — началу 1960-х годов. К этому же времени в общих чертах были выяснены функции и принципы структурной организации РНК. Были от-

крыты рибосомы и рибосомные РНК (рРНК), транспортные РНК (тРНК) и, наконец, информационные (или «мессенджер», мРНК). Стало ясным, что в совокупности все эти РНК служат промежуточным звеном при переносе генетической информации от ДНК к белкам. Было доказано, что собственно биосинтез полипептидных цепей белка происходит на рибосомах, где генетическая информация, переписанная (транскрибированная) с ДНК в виде мРНК, транслируется с помощью тРНК. В 1959—1961 гг. П. Доти, А. С. Спириным и их сотр. было установлено, что все клеточные РНК представляют собой однотяжевые полинуклеотиды с характерной вторичной структурой, построенной из коротких двуспиральных участков, соединенных однотяжевymi участками. Несколько позже в лабораториях Дж. Уотсона, А. С. Спирина и М. Номуры были установлены принципы структурной организации рибосом.

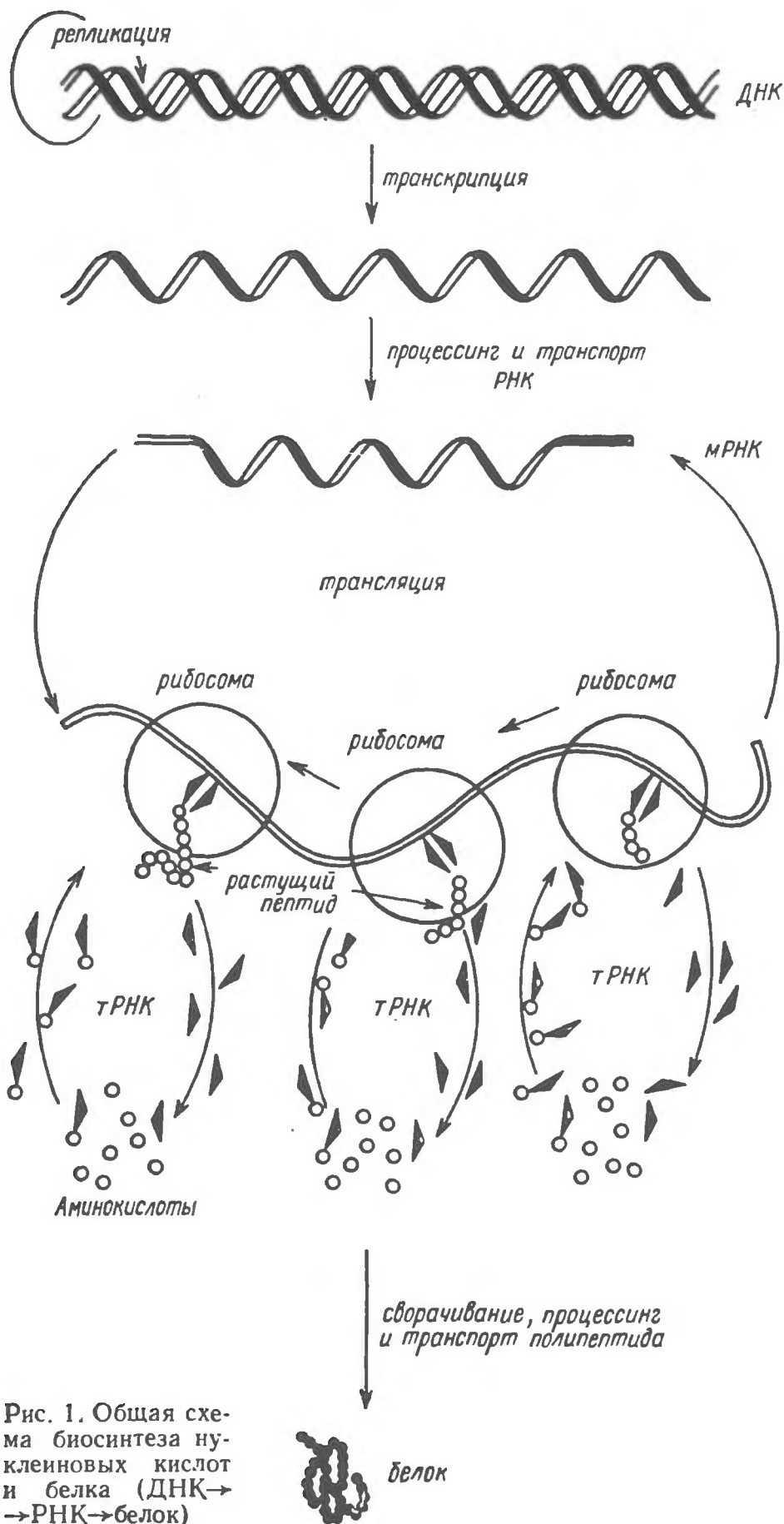
В 1960 г. сразу в нескольких лабораториях был открыт фермент РНК-полимераза, осуществляющий синтез РНК на ДНК-матрицах. Таким образом, идея Ф. Крика о передаче генетической информации от ДНК к белку через РНК (ДНК — РНК — белок), высказанная им еще в середине 1950-х годов, приобрела конкретное содержание (рис. 1).

Генетический аминокислотный код был полностью расшифрован в 1961—1966 гг. усилиями лабораторий М. Ниренберга, С. Очоа и Г. Кораны.

Открытие основных компонентов систем транскрипции и трансляции послужило важным стимулом в изучении механизмов регуляции этих процессов. В 1961 г. Ф. Жакоб и Ж. Моно опубликовали схему регуляции синтеза белков на уровне транскрипции при помощи регуляторных белков, а в 1966 г. У. Гилберт и Б. Мюллер-Хилл впервые выделили такой белок. Кроме того, оказалось, что РНК-полимераза сама является регулятором генной активности (Р. Б. Хесин, 1962—1966). Эти работы привели к открытию основных регуляторных генетических элементов — промоторов и терминаторов транскрипции.

В середине 1960-х годов начались исследования нуклеотидных последовательностей РНК. Первыми были определены первичные структуры тРНК (Р. Холли и сотр., 1965; А. А. Баев и сотр., 1967). Развитие техники фракционирования фрагментов нуклеиновых кислот и прежде всего гель-электрофореза (Ф. Сэнгер и сотр.) позволило в начале 1970-х годов приступить к изучению первичной структуры высокомолекулярных РНК. В 1976—1978 гг. были созданы исключительно быстрые и эффективные методы секвенирования ДНК и РНК (А. Максам и У. Гилберт, Ф. Сэнгер и сотр.), которые позволили за короткое время получить огромную информацию о первичной структуре генов, их регуляторных элементах, вирусных и рибосомных РНК и т. д.

В 1973 г. одновременно в лабораториях А. Рича и А. Клуга с помощью рентгеноструктурного анализа была установлена про-



странственная структура тРНК. В эти же годы был открыт основной структурный элемент эукариотических хромосом — нуклеосома — и разработаны методы ее исследования.

В 1970 г. Г. Темин и Д. Балтимор открыли в онкогенных вирусах РНК-зависимую ДНК-полимеразу и тем самым показали, что в принципе поток генетической информации может быть обращен от РНК к ДНК.

Огромное значение для молекулярной биологии последнего десятилетия имеет развитие генетической инженерии (возникшей в 1972—1973 гг.; П. Берг, П. Лобан, С. Коэн и Г. Бойер) и методов работы с рекомбинантными ДНК в сочетании с методами химического синтеза крупных фрагментов ДНК. В результате сделались доступными для исследования индивидуальные гены и регуляторные генетические элементы, было стимулировано изучение ферментов биосинтеза и обмена нуклеиновых кислот. Благодаря этому после 1977 г. были обнаружены мозаичное (экзон-интронное) строение генов, явление сплайсинга и ферментативной активности у РНК, усилители («энхансеры») экспрессии генов, многие регуляторные белки, онкогены и онкобелки, мобильные генетические элементы. Возникла белковая инженерия, которая позволяет получать новые, не существующие в природе белки. Молекулярная биология начала оказывать существенное влияние на развитие биотехнологии, медицины и сельского хозяйства.



Вводная

Глава I

СТРУКТУРА НУКЛЕИНОВЫХ КИСЛОТ

У эукариотических организмов ДНК локализована преимущественно в ядрах клеток; у прокариот она образует довольно компактный нуклеоид, в котором содержится вся хромосома бактериальной клетки. Такие клеточные органеллы, как митохондрии и хлоропласты, имеют свою собственную ДНК. Кроме того, в цитоплазме многих прокариот и низших эукариот обнаруживаются внехромосомные ДНК — плазмиды.

Рибосомные, транспортные и информационные РНК локализованы и функционируют в цитоплазме про- и эукариотических клеток. У эукариот они синтезируются в клеточном ядре, где и обнаруживаются их предшественники. Кроме того, в ядрах и цитоплазме клеток имеется множество так называемых малых РНК.

Всякому структурному исследованию ДНК или РНК предшествуют выделение их из клеток, очистка и фракционирование. Поскольку в клетке нуклеиновые кислоты практически всегда находятся в комплексе с белками (т. е. в виде *нуклеопротеидов*), их выделение сводится в основном к очистке от белков (депротеинизации). Чаще всего нуклеиновые кислоты экстрагируют из гомогенатов клеток или очищенных клеточных органелл смесью фенол — вода в присутствии ионных детергентов (например, додецилсульфата натрия). При этом белки (и ряд других клеточных компонентов) переходят в органическую фазу, а нуклеиновая кислота остается в водной фазе. Из водного раствора ДНК или РНК осаждают спиртом.

Среди множества методов фракционирования нуклеиновых кислот отметим здесь электрофорез в гелях, высокая разрешающая способность которого позволяет получить любую ДНК или РНК в индивидуальном состоянии.

1. ПЕРВИЧНАЯ СТРУКТУРА НУКЛЕИНОВЫХ КИСЛОТ

Нуклеиновые кислоты относятся к классу биологических полимеров. Поэтому в понятие *первичная структура* ДНК и РНК, как и любого другого биополимера, входит строение их мономерных ос-

татков, химическая природа межмономерных ковалентных связей и последовательность чередования мономерных звеньев в полимерной цепи.

Нуклеозиды и нуклеотиды

Мономерными звеньями ДНК и РНК являются остатки нуклеотидов. *Нуклеотиды* — это фосфорные эфиры *нуклеозидов*, которые, в свою очередь, построены из остатка углевода — пентозы и *гетероциклического основания*. В РНК углеводные остатки представлены *D-рибозой*, в ДНК — *2-D-дезоксирибозой*. Связь между углеводным остатком и гетероциклическим основанием в нуклеозиде осуществляется через атом азота в основании, т. е. с помощью *N-гликозидной связи*. Таким образом, нуклеозидные остатки в ДНК и РНК относятся к классу *N-гликозидов*. Как уже отмечалось во Введении, в качестве гетероциклических оснований ДНК содержат два пурина: *аденин* и *гуанин* — и два пиримидина: *тимин* и *цитозин*. В РНК вместо тимина содержится *урацил*. Кроме того, ДНК и РНК обычно содержат так называемые минорные нуклеотидные остатки — производные обычных нуклеотидов по основаниям или углеводному остатку, доля которых в зависимости от вида нуклеиновой кислоты колеблется от десятых процента до десятков процентов. Строение, химическая номенклатура и принятые сейчас сокращенные обозначения нуклеотидов и их компонентов показаны на рис. 2.

Межнуклеотидная связь. Нуклеазы

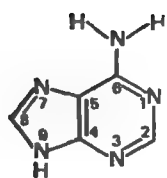
В нуклеиновых кислотах мономерные остатки связаны между собой *фосфодиэфирными связями*. Как в ДНК, так и в РНК эта связь осуществляется только за счет *3'-ОН* одного нуклеозидного остатка и *5'-ОН* другого (рис. 3). Поэтому межнуклеотидную связь называют *3'—5'-фосфодиэфирной*. Отсюда следует, что полинуклеотидные цепи ДНК и РНК полярны и их концевые остатки неравноценны: если рассматривать нефосфорилированный по концам полинуклеотид, то на одном конце он будет содержать *5'-*, а на другом конце — *3'-ОН*. Эти концы называют *5'-* и *3'-концами цепи*, соответственно.

Из характера межнуклеотидной связи следует также, что нуклеиновые кислоты — это неразветвленные полимеры. Лишь в очень редких случаях (они относятся, как правило, к промежуточным продуктам биосинтеза РНК) *2'-гидроксильная группа* нуклеотидных остатков может использоваться для образования дополнительной *фосфодиэфирной связи* (см. гл. VIII, раздел 4).

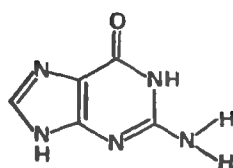
На свойствах фосфодиэфирных связей сильно сказывается наличие или отсутствие *ОН-группы* при *C2'-атоме* рибозы. Так, меж-

Основания

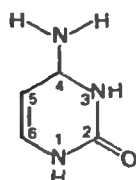
Пурины



аденин (A)

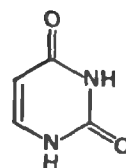


гуанин (G)

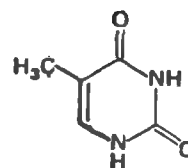


цитозин (C)

Пиримидины

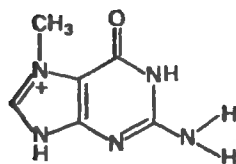


урацил (U)

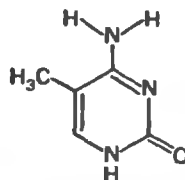


тимин (T)

Минорные основания

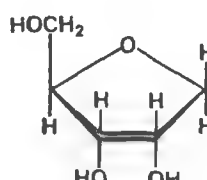


7-метилгуанин
(7 MeG)

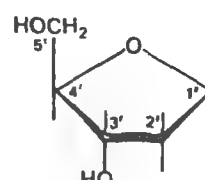


5-метилцитозин
(5 MeC)

Сахара

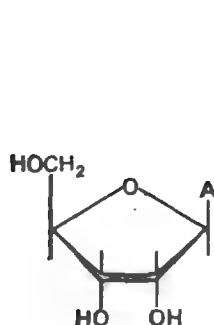


D-рибоза

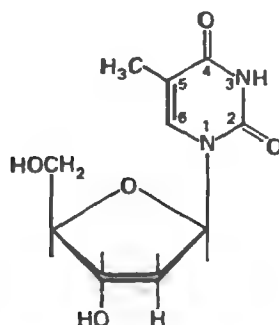


D-дезоксирибоза

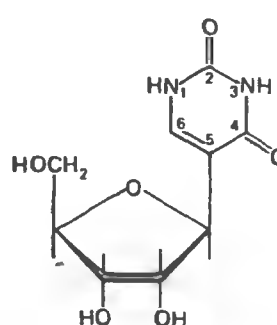
Нуклеозиды



аденозин (rA)



тимидин (dT)



псевдоуридин (ψ)

Минорный нуклеозид

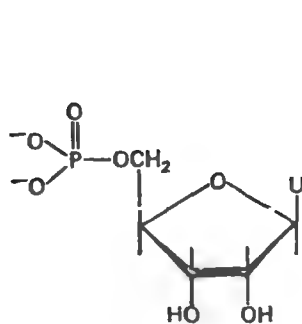
Другие рибонуклеозиды

гуанозин (rG)
цитидин (rC)
уридин (rU)

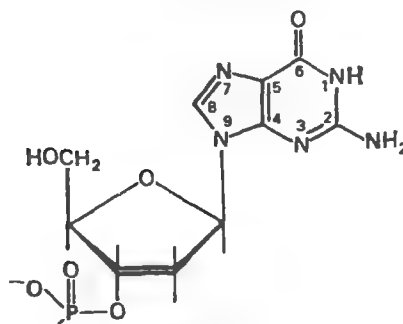
Другие дезокси- рибонуклеозиды

дезоксигуанозин (dG)
дезоксиаденозин (dA)
дезоксцитидин (dC)

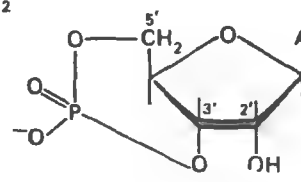
Нуклеотиды и их производные



уридин-5'-фосфат
(rU)



дезоксигуанозин-3'-фосфат
(dGp)



аденозин-3',5'-циклофосфат
(с AMP)

Другие рибонуклеотиды

аденозин-3'-фосфат (Ap)
цитидин-5'-фосфат (pC)
гуанозин-5'-фосфат (pG)

Другие дезоксирибонуклеотиды

дезоксиаденозин-5'-фосфат (pdA)
дезоксцитидин-3'-фосфат (dCp)
тимидин-5'-фосфат (pT)

Рис. 2. Структурные формулы, номенклатура и сокращенные обозначения компонентов нуклеиновых кислот

нуклеотидные связи в РНК значительно лабильнее, чем в ДНК; они легко гидролизуются в присутствии щелочи.

В клетке межнуклеотидные связи в ДНК и РНК расщепляются *нуклеазами* — обширным классом ферментов, представители которого различаются по механизму действия и специфичности (табл. 1). Среди нуклеаз, приведенных в таблице, нужно особо выделить *эндонуклеазы рестрикции (рестриктазы)*. Эти ферменты (их функции рассмотрены в гл. VI) узнают в молекулах ДНК не отдельные нуклеотидные остатки, а нуклеотидные последовательности из четырех, пяти или шести остатков и поэтому расщепляют любую ДНК на сравнительно небольшое число строго определенных фрагментов.

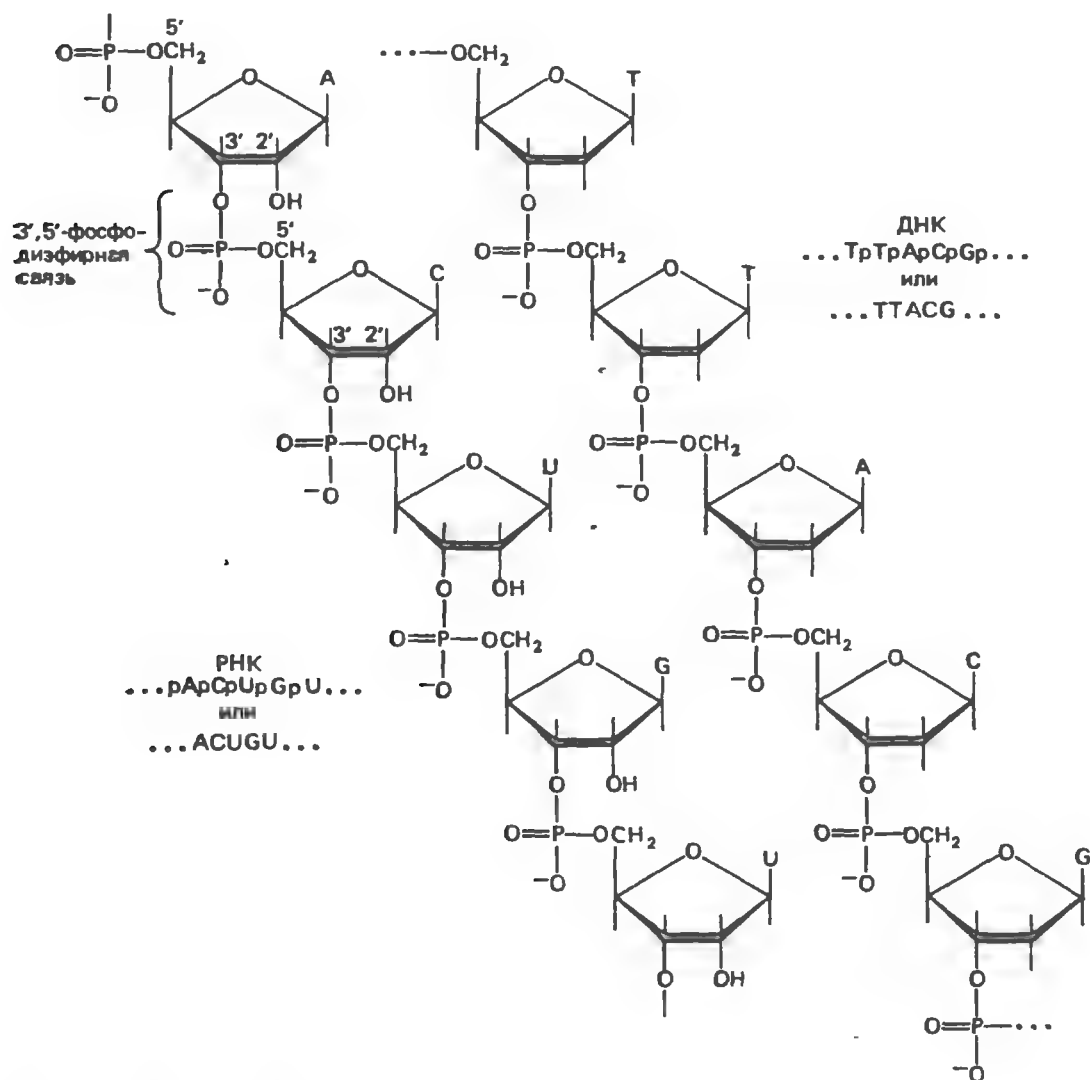


Рис. 3. Полинуклеотидные цепи РНК и ДНК

Обратите внимание на то, как сокращению записываются их нуклеотидные последовательности

Таблица 1. Примеры нуклеаз, используемых для изучения структуры нуклеиновых кислот

Фермент	Тип нуклеазной активности	Субстрат	Продукты реакции
Нуклеаза S_1 из <i>Aspergillus oryzae</i>	Неспецифическая эндонуклеаза	Однотяжевые РНК и ДНК	5'-фосфорилированные олигонуклеотиды
Фосфодиэстераза из яда змей	Неспецифическая 3' → 5'-экзонуклеаза	Олиго- и полинуклеотиды со свободным 3'-концом	Нуклеозид-5'-фосфаты
Эксонуклеаза III из <i>E. coli</i>	Неспецифическая ДНКаза, 3' → 5'-эксонуклеаза	Двухтяжевая ДНК	То же
Гуаниловая РНКаза T_1 из актиномицетов	Специфическая РНКаза, эндонуклеаза	Межнуклеотидные связи в однотяжевых РНК, образованные 3'-фосфатным остатком гуанозина и 5'-гидроксилом любого нуклеотида	Олигонуклеотиды с остатком гуанозин-3'-фосфата на 3'-конце
РНКаза V_1 из яда кобры	Неспецифическая эндонуклеаза	Двухтяжевые участки РНК	5'-фосфорилированные олигонуклеотиды
РНКаза III из <i>E. coli</i>	То же	То же	3'-фосфорилированные олигонуклеотиды
ДНКаза I из поджелудочной железы	»	ДНК (предпочтительно двухтяжевая)	То же
Эндонуклеаза (рестриктаза) <i>EcoRI</i> из <i>E. coli</i>	Специфическая ДНКаза	Последовательность GAATTC CTTAAG ↑ в двухтяжевой ДНК *	Специфический набор двухтяжевых фрагментов ДНК, фосфорилированных по 5'-концам
Эндонуклеаза (рестриктаза) <i>AluI</i> из <i>Arthrobacter luteus</i>	То же	Последовательность AGCT TCGA ↑ в двухтяжевой ДНК *	То же

* Стрелками показаны расщепляемые межнуклеотидные связи.

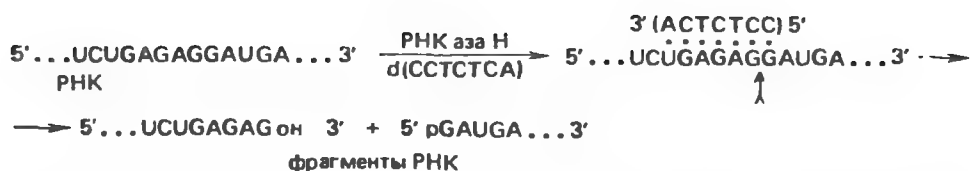


Рис. 4. Направленное расщепление полинуклеотидной цепи РНК ферментом РНКазой Н

В качестве примера показана фрагментация цепи рибосомной 16S РНК *Escherichia coli*

Рестриктазы незаменимы в структурных исследованиях нуклеиновых кислот. В случае РНК достаточно специфическое расщепление ее цепи можно осуществить с помощью рибонуклеазы Н, гидролизующей полирибонуклеотиды только в ДНК — РНК-гибридах. Для этого с участком РНК, предназначенным для расщепления, предварительно связывают комплементарный ему олигодезоксирибонуклеотид и обрабатывают образовавшийся комплекс РНК-азой Н (рис. 4).

Определение нуклеотидной последовательности ДНК

Размеры двуспиральных ДНК характеризуют числом пар нуклеотидов (п. н.), приходящихся на одну макромолекулу. Для клеточных и вирусных ДНК они варьируют в очень широких пределах. Так, например, наиболее изученные бактериальные плазмиды и ДНК многих вирусов и бактериофагов содержат несколько тысяч пар нуклеотидов (т. п. н.), ДНК половых факторов бактерий, митохондрий и хлоропластов — несколько десятков или сотен т. п. н. Размеры хромосом бактерий — несколько миллионов п. н., дрожжей — порядка 10^7 п. н. Суммарная длина хромосомных ДНК человека составляет около $3 \cdot 10^9$ п. н.

Современные методы определения нуклеотидной последовательности ДНК позволяют в одном эксперименте просеквенировать (от англ. sequence — последовательность) ее фрагмент длиной в 150—300 нуклеотидных остатков (н. о.). Поэтому исходная молекула ДНК предварительно фрагментируется. Для этого чаще всего ДНК гидролизуют рестриктазами (причем проводится независимое расщепление двумя рестриктазами или более, в результате чего образуются перекрывающиеся фрагменты (рис. 5). Это позволяет после определения нуклеотидной последовательности соответствующих фрагментов реконструировать первичную структуру всей ДНК.

Существует два принципиально различных подхода к определению нуклеотидной последовательности ДНК. Первый из них предложен А. Максамом и У. Гилбертом и основан на специфическом химическом расщеплении полинуклеотидной цепи. Последовательность операций при секвенировании ДНК методом Максама —

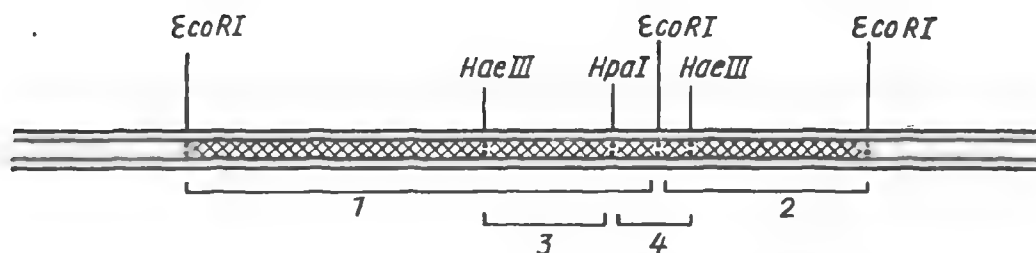


Рис. 5. Расщепление рестриктазами крупного сегмента гена β -цепи глобина кролика с образованием перекрывающихся фрагментов

Секвенирование фрагмента 4, получаемого совместным действием рестриктаз *Hae III* и *Hpa I*, позволяет однозначно расположить друг относительно друга фрагменты 1 и 2

Гилберта в общих чертах показана на рис. 6. Рассмотрим здесь основные принципы этого метода.

1. Фрагмент ДНК должен быть гомогенен по тому из концов, по которому вводится радиоактивная метка.

2. Образец анализируемой цепи ДНК делится на четыре равных порции, и условия химической модификации каждой из них подбираются таким образом, чтобы в реакции участвовало либо только одно, либо преимущественно одно из двух пуриновых или пиримидиновых оснований; пуриновые основания метилируют по N7-атому диметилсульфатом, пиримидиновые — расщепляют обработкой гидразином; обработка химическими реагентами приводит в обоих случаях к появлению модифицированного мономерного звена, которое может быть легко выщеплено из полинуклеотидной цепи.

3. Химическая модификация проводится в таких ограничивающих условиях, когда в каждой молекуле ДНК модифицируется только одно из звеньев данного типа, причем с равной вероятностью независимо от его положения в цепи; если это условие выполняется, то в итоге в каждой порции получают популяцию фрагментов ДНК с единственным модифицированным нуклеотидным остатком, но в сумме все остатки данного вида (G, например) будут промодифицированы.

4. Расщепление цепи сопровождается выщеплением модифицированного звена; если теперь измерить длину образовавшегося фрагмента с точностью до одного нуклеотидного остатка, то можно определить положение модифицированного остатка; совокупность длин фрагментов для случая модификации одного из четырех типов оснований даст нам, таким образом, информацию о распределении этого основания вдоль цепи исходной ДНК.

5. Длины фрагментов определяют с помощью высокоразрешающего электрофореза в тонком слое полиакриламидного геля. Электрофорез проводится при повышенной температуре в буферной системе, содержащей 7 М мочевину, что способствует разрушению вторичной структуры фрагментов; детекция зон фрагментов на электрофореграмме зависит от способа мечения его конца (если это

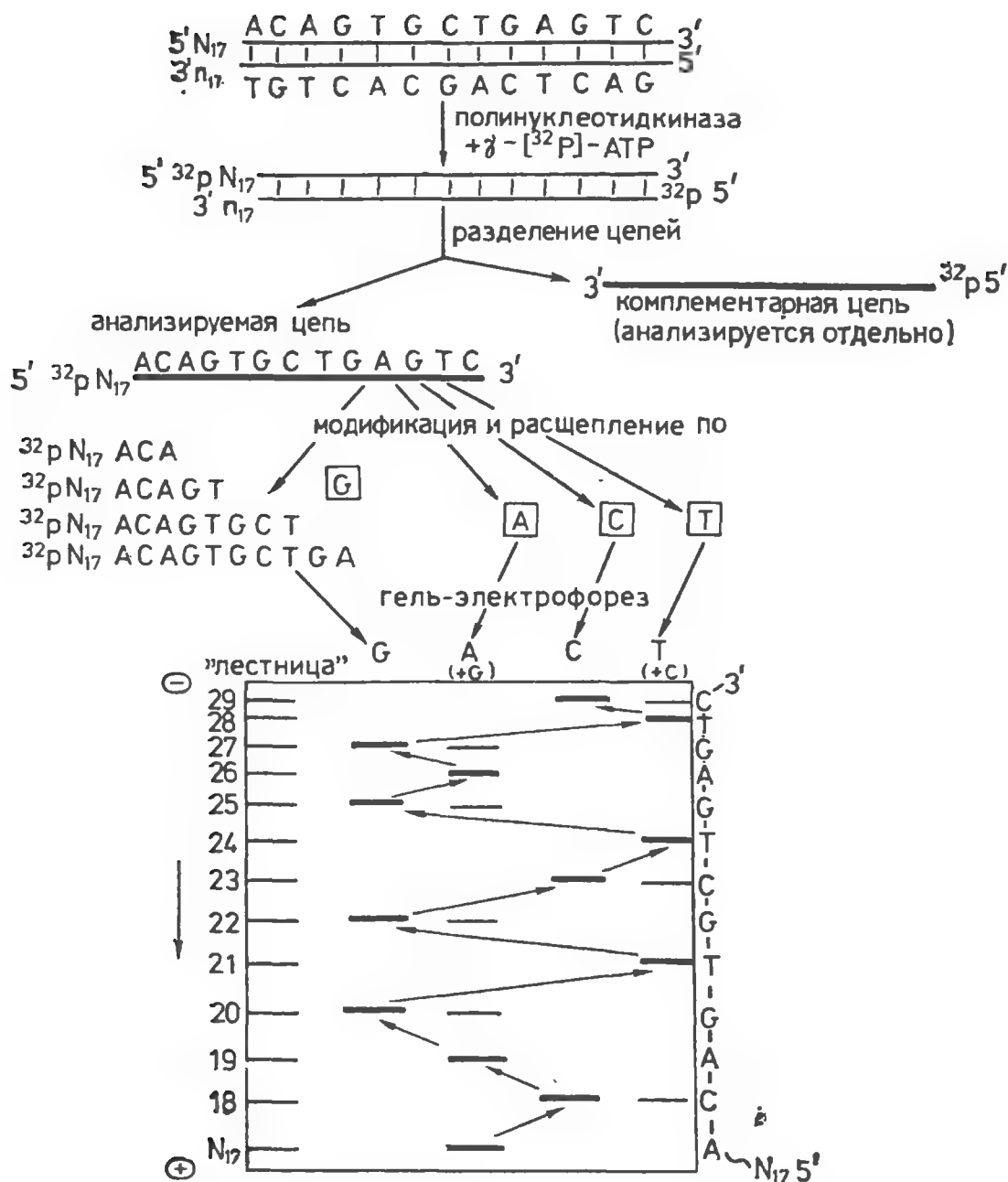


Рис. 6. Схема определения 13-членной нуклеотидной последовательности в 30-членном фрагменте ДНК по Максому — Гилберту

Обратите внимание на то, что медленнее всего при гельэлектрофорезе движется 29-звенный фрагмент (образующийся при отщеплении 3'-концевого остатка C), а быстрее всего — 17-звенный фрагмент

радиоактивная метка на 5'-конце, как на рис. 6, то применяется радиоавтография; при этом, естественно, детектируются только 5'-концевые фрагменты); собственно измерение длины фрагмента осуществляется сравнением его подвижности со стандартом («лестница» на рис. 6).

Второй подход разработан Ф. Сэнгером и сотр. В его основе лежит копирование анализируемого участка ДНК с помощью

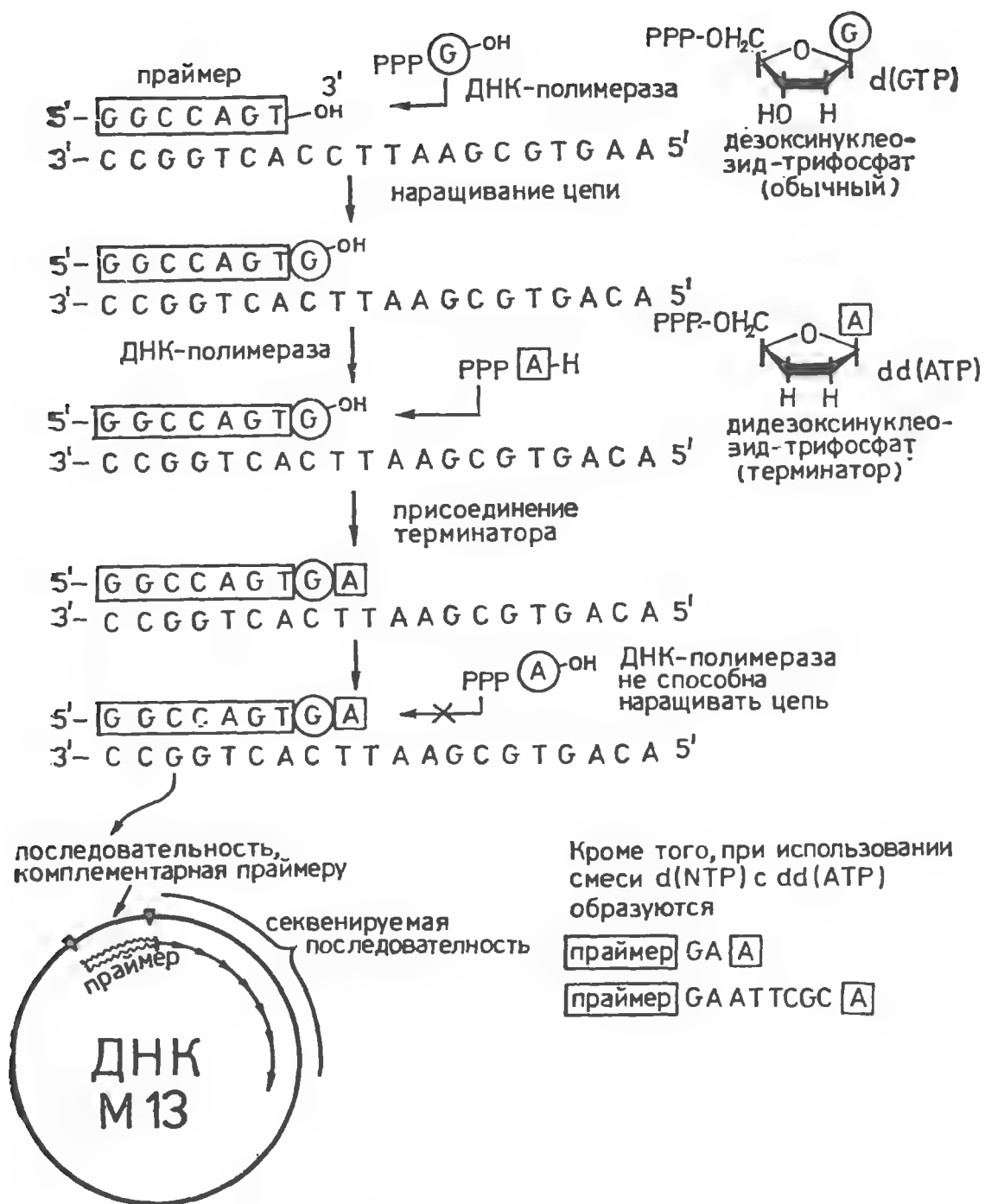
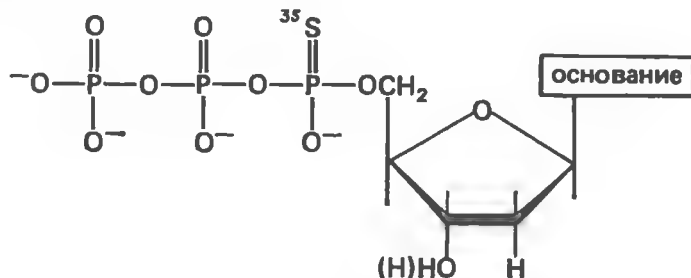


Рис. 7. Схема, иллюстрирующая принципы секвенирования ДНК по Сэнгеру. Показана терминация только по остаткам А.

ДНК-полимеразной реакции (рассматриваемой в гл. II при обсуждении репликации ДНК). Принцип метода показан на рис. 7. Как и при секвенировании ДНК по Максому — Гилберту, здесь добиваются, чтобы обрыв цепи (при попадании на ее 3'-конец комплементарного *терминаторного* нуклеотидного остатка) происходил равновероятно по всем положениям данного остатка (рис. 7 — остаток Т). Это достигается экспериментальным выбором «правильного» соотношения дезоксинуклеозидтрифосфатов и терминаторного

дидезоксинуклеозидтрифосфата. В итоге в каждой из четырех реакционных смесей образуются примерно в равных количествах молекулы комплементарных полинуклеотидов, синтез которых остановлен на анализируемом нуклеотидом остатке.

Для секвенирования ДНК по Сэнгеру созданы универсальные конструкции (на базе генома фага M13) для встраивания в них изучаемого фрагмента *, позволяющие применять универсальный праймер. Метка вводится в праймер или по α -фосфатной группе нуклеозидтрифосфата. В последнем случае это может быть не только меченый атом фосфора (^{32}P , например), но и атом серы (^{35}S), заменяющий атом кислорода:



Меченые фрагменты разделяют так же, как и при секвенировании ДНК по Максаму — Гилберту.

В настоящее время существует множество вариантов как метода Максама — Гилберта, так и метода Сэнгера. Главное, эти методы удалось полностью автоматизировать. Так, например, при секвенировании ДНК по Сэнгеру на 5'-конец праймера вводят флуоресцентные метки, причем для каждого из четырех анализируемых нуклеотидов используются флуоресцирующие агенты с различными спектральными характеристиками. После электрофоретического разделения гель сканируется при четырех различных длинах волн и полученная информация сразу обрабатывается на ЭВМ. При этом все биохимические операции также проводятся роботом.

Применение автоматов позволяет в принципе определять полную нуклеотидную последовательность ДНК, составляющей геном любого организма, включая геном человека.

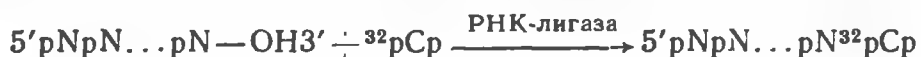
Определение нуклеотидной последовательности РНК

Размеры полинуклеотидных цепей РНК колеблются от нескольких десятков и сотен (транспортные, малые рибосомные, малые ядерные РНК) до нескольких тысяч (рибосомные, информационные и вирусные РНК) нуклеотидных остатков.

Таким образом, полная нуклеотидная последовательность многих классов РНК (длиной до 250—300 н. о.) может быть прямо оп-

* Методы получения рекомбинантных ДНК в этой книге не рассматриваются. Читателям рекомендуется обратиться к руководствам по генетической инженерии, приведенным в конце этой главы.

ределена с помощью одного из вариантов метода Максама — Гилберта. Как и в случае фрагментов ДНК, на один из концов РНК вводится радиоактивная метка. Наряду с фосфорилированием 5'-конца полинуклеотидной цепи (с помощью полинуклеотидкиназы и γ -[^{32}P]-АТР) здесь применяется введение радиоактивной метки на 3'-конец молекулы. Для этого к 3'-концевой ОН-группе РНК пришивают меченый по 5'-фосфатной группе цитидин — 3', 5'-дифосфат (5'-[^{32}P]-pCp) с помощью фермента РНК-лигазы:



Наряду с химическим расщеплением секвенирование РНК можно проводить путем гидролиза ее цепи специфическими РНК-азами.

Часто для определения нуклеотидной последовательности достаточно протяженной полинуклеотидной цепи РНК наиболее простым оказывается секвенирование гена (т. е. ДНК), в котором эта РНК закодирована. В тех случаях, когда это сделать затруднительно, применяется секвенирование РНК по Сэнгеру с использованием обратной транскриптазы вместо обычной ДНК-полимеразы. Этот фермент, закодированный в геноме многих вирусов (см. гл. 13, раздел 3), синтезирует на РНК-матрице комплементарную ей полидезоксирибонуклеотидную цепь. Для его работы требуется *затравка* (*праймер*) со свободным 3'-концом, в качестве которой используется олигонуклеотид, комплементарный 3'-концевой части РНК. Для полиаденилированных РНК (см. гл. VIII, раздел 4) затравкой служит олиготимидиловая кислота. В других же случаях требуется дополнительная информация о первичной структуре 3'-концевого участка РНК, на основании которой химически синтезируют олигонуклеотид — затравку. Само же секвенирование проводится так, как показано на рис. 7.

2. МАКРОМОЛЕКУЛЯРНАЯ СТРУКТУРА НУКЛЕИНОВЫХ КИСЛОТ

Пространственная организация макромолекул ДНК и РНК задается их нуклеотидной последовательностью и описывается структурами двух уровней: вторичной и третичной.

Вторичная структура нуклеиновых кислот создается за счет взаимодействия соседних по полинуклеотидной цепи мономерных звеньев, а в случае двуспиральных молекул (или участков молекул) также взаимодействием нуклеотидных остатков, находящихся напротив друг друга в двойной спирали. *Третичная структура* нуклеиновых кислот организуется за счет взаимодействия нуклеотидных остатков, принадлежащих различным элементам их вторичной структуры.

Макромолекулярная структура ДНК

а. Двойная спираль ДНК. Модель Уотсона и Крика

В 1953 г. Дж. Уотсон и Ф. Крик, опираясь на известные данные о структуре мономерных остатков нуклеиновых кислот и открытое Э. Чаргаффом правило, согласно которому в любой ДНК количество аденина равно количеству тимина, а количество гуанина — количеству цитозина ($A=T$, $G=C$), расшифровали рентгенограммы псевдокристаллической формы (ориентированных волокон) ДНК. Согласно их модели, молекула ДНК представляет собой правильную спираль, образованную двумя полинуклеотидными цепями, закрученными друг относительно друга и вокруг общей оси. Диаметр спирали постоянен вдоль всей ее длины и равен

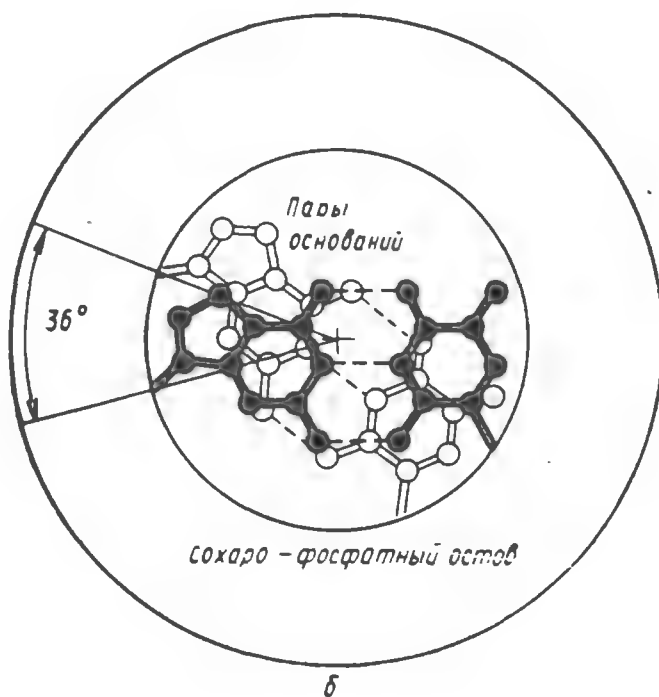
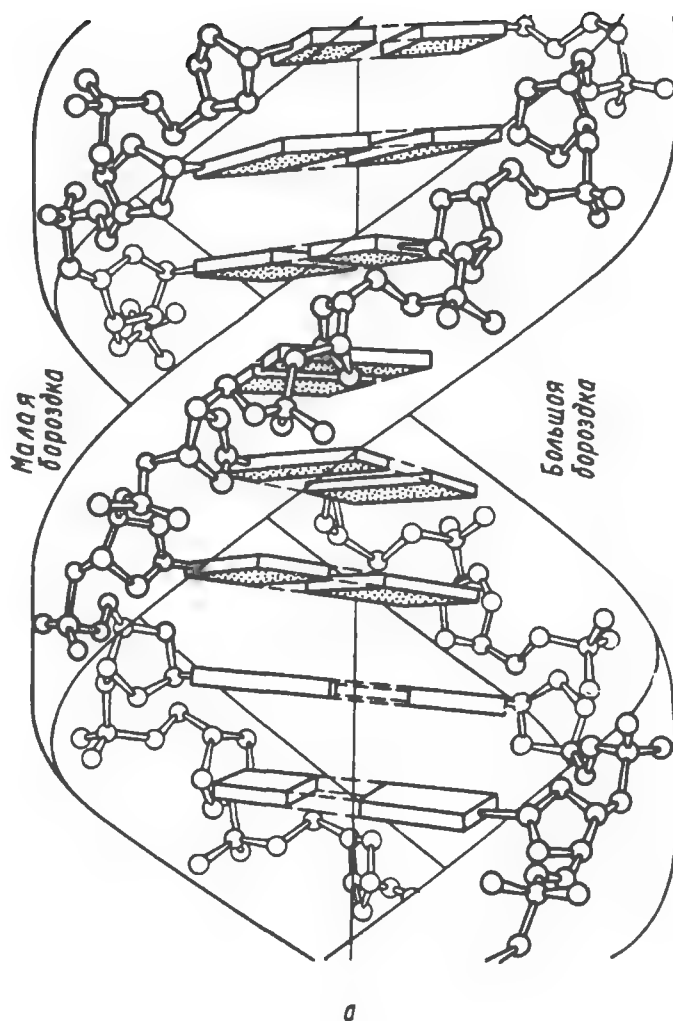


Рис. 8. В-форма двойной спирали ДНК:

а — общий вид спирали (длинная ось спирали лежит в плоскости рисунка) (хорошо виден сахарно-фосфатный остов молекулы, обращенный наружу). Комплементарные пары схематически изображены прямоугольниками, плоскости которых почти перпендикулярны оси спирали); б — вид сверху (показаны две комплементарные пары оснований. Хорошо видно, что основания заполняют центральную часть спирали)

1,80 нм. Длина витка спирали, который соответствует ее периоду идентичности, составляет примерно 3,40 нм. На один виток спирали приходится 10 нуклеотидных остатков в одной цепи. Таким образом межнуклеотидное расстояние вдоль оси спирали составляет 0,34 нм (рис. 8).

Для поддержания регулярности спирали необходимо, чтобы против остатка пуринового основания в одной цепи находился остаток пиримидинового основания в другой. Дж. Уотсон и Ф. Крик нашли, что это требование выполняется только в том случае, если А образует стабилизированную водородными связями пару с Т, а Г с С (см. следующий раздел).

Комплементарные основания в двойной спирали ДНК лежат в одной плоскости, причем эта плоскость практически перпендикулярна главной оси спирали. Соседние пары оснований в ДНК повернуты друг относительно друга на 36° [угол между прямыми, соединяющими $C1'$ в соседних комплементарных парах (см. рис. 12)].

Из рис. 8 видно, что углеводофосфатный остов молекулы обращен наружу. Спираль закручена таким образом, что на ее поверхности можно выделить две бороздки: большую шириной 2,20 нм и малую шириной около 1,20 нм (их называют также главным и минорным желобками). Спираль — правозакрученная, а полинуклеотидные цепи в ней антипараллельны. Это означает, что, если двигаться вдоль оси спирали от одного ее конца к другому, в одной цепи будут встречаться фосфодиэфирные связи в направлении $3' \rightarrow 5'$, а в другой — в направлении $5' \rightarrow 3'$. Иными словами, на каждом из концов линейной молекулы ДНК расположены $5'$ -конец одной и $3'$ -конец другой цепи.

Позднее было установлено, что модель Уотсона и Крика описывает структуру одной из нескольких форм двойной спирали, названной *В-формой*. Это основная форма двуспиральной ДНК, в которой большая часть ее молекул существует в клетке.

Для того чтобы более подробно описать В-форму ДНК, а также превращения ее в другие формы двойной спирали, рассмотрим пространственную структуру мономерных звеньев нуклеиновых кислот и межмолекулярные взаимодействия, приводящие к образованию двойной спирали.

б. Конформация нуклеотидных остатков и межнуклеотидные взаимодействия в нуклеиновых кислотах

Все пять гетероциклических оснований, входящих в состав нуклеиновых кислот, имеют плоскую конформацию. В то же время для остатков рибозы и дезоксирибозы плоская конформация (когда атомы углерода $C1'$, $C2'$, $C3'$, $C4'$ и гетероатом кислорода находятся в одной плоскости) энергетически невыгодна. Среди многочисленных теоретически возможных конформаций этих остатков в полинуклеотидах реализуются только две: либо $C2'$ -эндо-, либо $C3'$ -

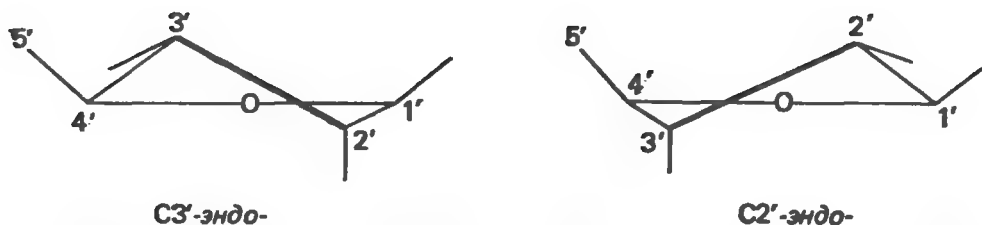


Рис. 9. Две конформации углеводных остатков, встречающиеся в ДНК и РНК:

$C1'$, O и $C4'$ расположены в одной плоскости, $C2'$ или $C3'$ находятся в *эндо*-конформации, когда они выведены над этой плоскостью (т. е. в направлении связи $C4' \rightarrow C5'$). Обратите внимание, что соседний $C3'$ - или $C2'$ -атом слегка выведен под плоскость (т. е. находится в слабой *экзо*конформации)

эндо-конформации (называемые также S- и N-конформациями соответственно; рис. 9).

Важнейший параметр пространственной структуры нуклеотидного звена — взаимное расположение остатков основания и углевода. Теоретический анализ конформации нуклеозидов и нуклеотидов и рентгеноструктурные исследования показали, что здесь существуют две области разрешенных конформаций, называемых *син*- и *анти*-конформациями. У пиримидиновых нуклеозидов в случае *анти*-конформации ближе всего к атому $C2'$ расположен атом водорода при $C6$, а в случае *син*-конформации — атом кислорода при $C2$. У пуриновых нуклеозидов в *анти*-конформации с атомом $C5'$ сближен атом водорода при $C8$, а в *син*-конформации атом $N3$ находится над плоскостью углеводного остатка (рис. 10).

В изолированных нуклеозидах и нуклеотидах переход от $C2'$ -*эндо*- к $C3'$ -*эндо*- и между *син*- и *анти*-конформациями происходит достаточно легко (хотя в зависимости от характера основания для данного соединения существуют предпочтительные конформации; так, например, пиримидиновые нуклеозиды находятся в основном в *анти*-конформации). Если же нуклеотидный остаток встроен в полинуклеотидную цепь, то его конформация при данных внешних условиях оказывается «замороженной».

Отметим также, что при изменении конформации углеводного остатка заметно изменяется расстояние между фосфатными группами в полинуклеотидной цепи (рис. 11).

Специфическая макромолекулярная структура нуклеиновых кис-

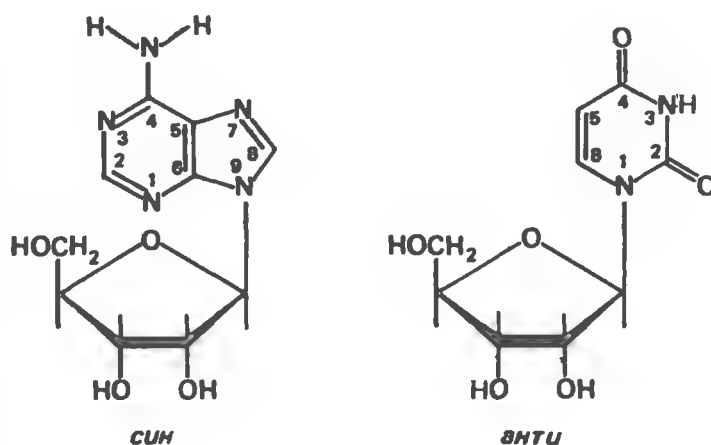


Рис. 10. *Син*- (аденозин) и *анти*- (уридин) конформации нуклеозидов

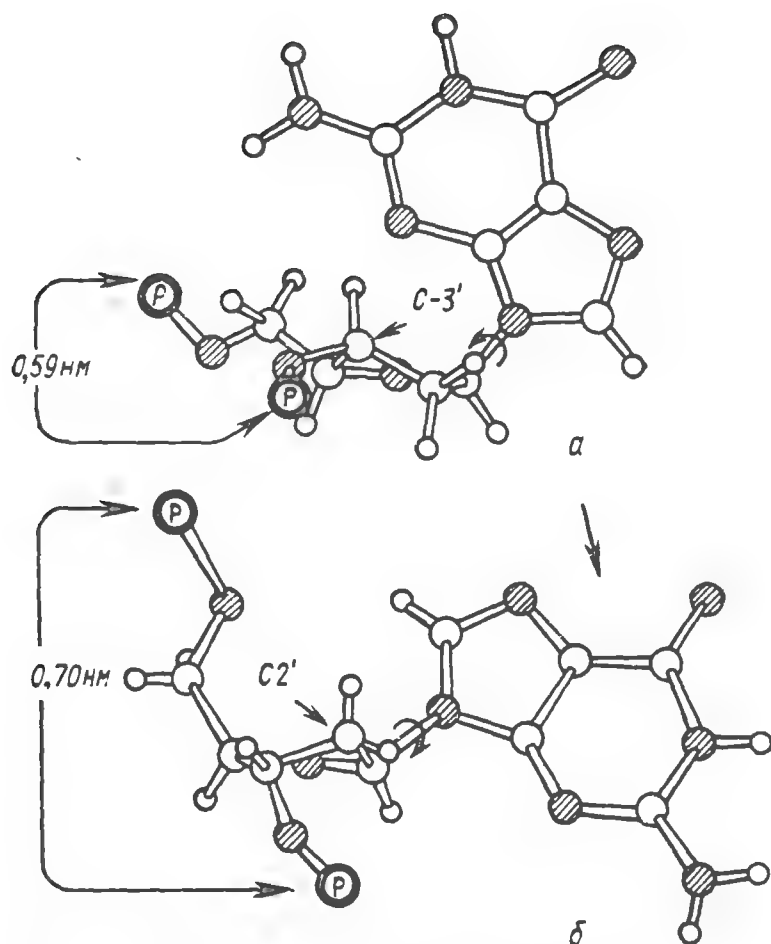


Рис. 11. Схема, показывающая изменение расстояний между фосфатными остатками нуклеотидного звена ДНК при $C3'$ -эндо- — $C2'$ -эндо-конформационном переходе:
 а — dG в *син*-конформации (Z форма); б — dG в *анти*-конформации (B-форма)

лот определяется двумя типами взаимодействий между гетероциклическими основаниями нуклеотидных остатков: 1) взаимодействиями между основаниями в комплементарных парах; 2) межплоскостными взаимодействиями оснований, расположенных друг над другом, или так называемыми «стэкинг»-взаимодействиями (от англ. *stacking* — укладываться в стопку).

Как уже отмечалось, образование комплементарных пар оснований (А·Т- и Г·С-пар) было впервые постулировано Дж. Уотсоном и Ф. Криком при создании модели двойной спирали ДНК. Строение этих пар показано на рис. 12. Обе пары близки по форме и имеют одинаковые размеры. Они связаны осью симметрии второго порядка, но при этом псевдосимметричны: при повороте на 180° вокруг оси, лежащей в плоскости рисунка (с выходом оснований из его плоскости), совпадают только $C1'$ -атомы.

Необходимо отметить, что кроме уотсон-криковских пар гетероциклические основания нуклеиновых кислот в принципе способ-

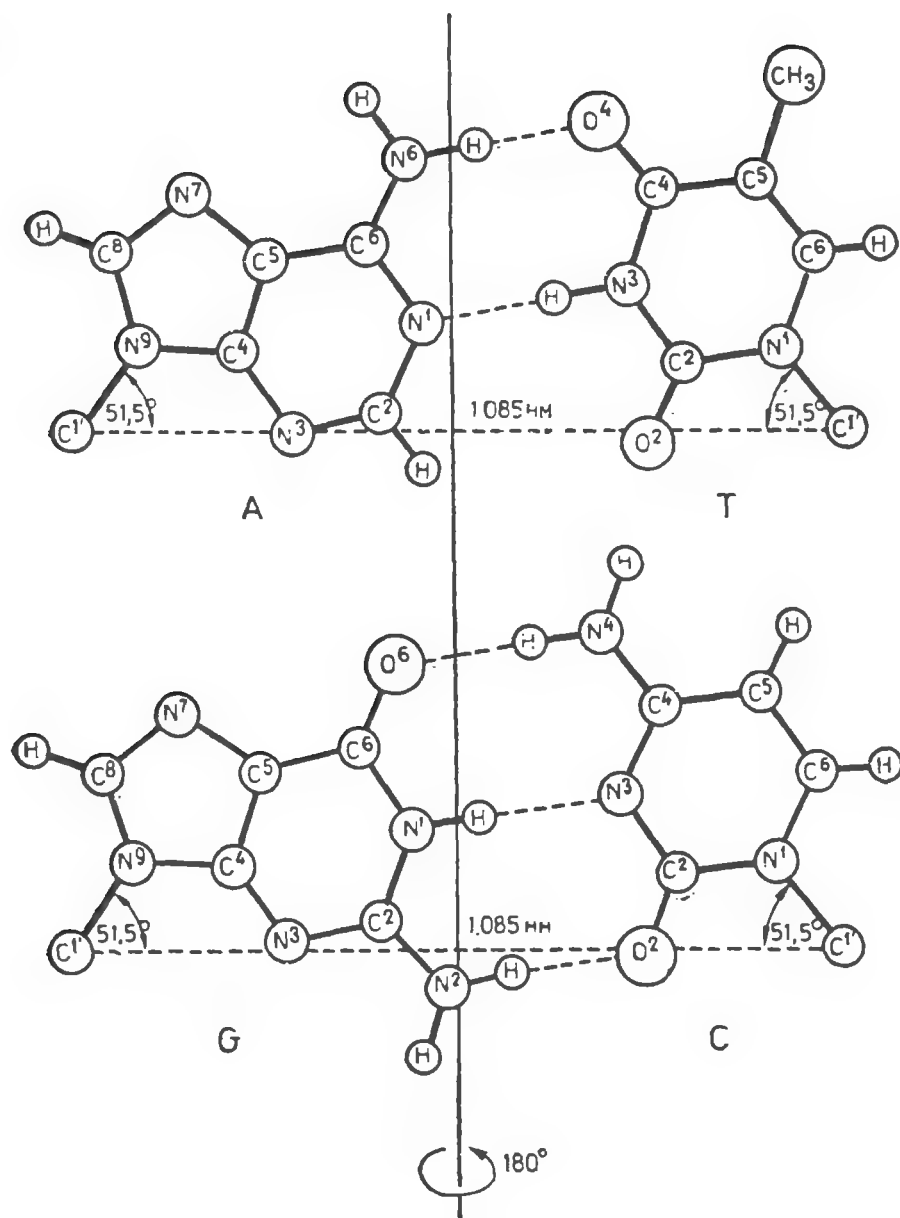


Рис. 12. Строение и некоторые параметры уотсон-криковских пар

ны образовывать множество связанных водородными связями пар другой структуры. Некоторые из этих пар обнаруживаются экспериментально для производных нуклеозидов и нуклеотидов, а также в комплексах ряда синтетических полинуклеотидов. Однако квантово-механические расчеты показывают, что уотсон-криковские А·Т-(в случае РНК — А·U-) и G·C-пары энергетически наиболее выгодны. Происходит это потому, что в этих парах центры с повышенной и пониженной электронной плотностью оснований расположены оптимально друг относительно друга. Таким образом, элементарные пары оснований в нуклеиновых кислотах стабилизированы преимущественно электростатическими взаимодействиями

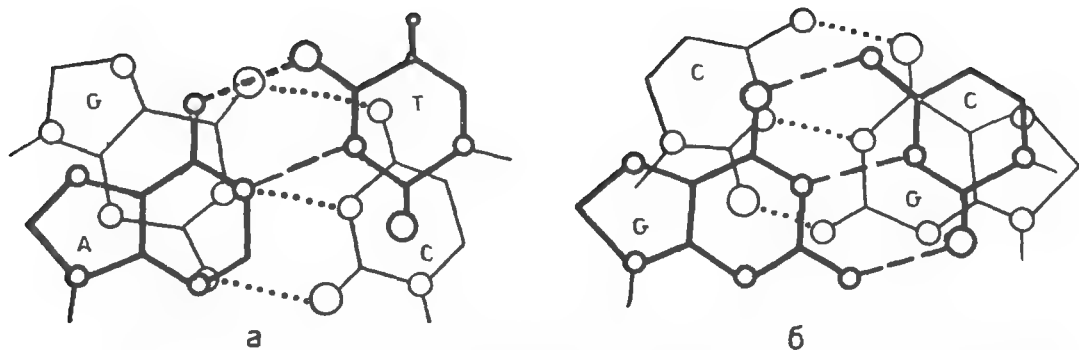


Рис. 13. Схема, иллюстрирующая перекрывание оснований в двойной спирали ДНК:

а — А·Т-пара находится сверху (жирные линии), Г·С-пара — снизу (тонкие линии); б — Г·С-пара — сверху; С·Г-пара — снизу; большие светлые кружки — атомы кислорода; малые — азота

(такую комплементарность называют электронной). Отметим также, что Г·С-пары существенно более стабильны, чем А·Т-пары.

Иную природу имеют межплоскостные взаимодействия оснований. Гетероциклические основания нуклеиновых кислот достаточно гидрофобны, т. е. в водном растворе им выгоднее расположиться друг над другом и тем самым уменьшить контакт с молекулами воды. При образовании таких стопок во взаимодействие вступают функциональные ($C=O$ и $C-NH_2$) группы одного основания и π -электронные системы соседнего с ним по вертикали основания. Поэтому стэкинг-взаимодействия оснований (в двойной спирали ДНК, например) зависят как от состава комплементарных пар, так и от их последовательности (рис. 13).

в. Полиморфизм двойной спирали ДНК

Уже на раннем этапе изучения пространственной структуры ДНК выяснилось, что при изменении внешних условий двойная спираль может принимать формы, отличные от уотсон-криковской (В-формы). Так, при понижении влажности (в препарате образца для рентгеноструктурного анализа) или активности воды в растворе (при добавлении спирта, например) ДНК переходит в так называемую *А-форму* (рис. 14). Этот переход связан в первую очередь с изменением конформации углеводного остатка. Если в В-форме ДНК остаток дезоксирибозы находится в $C2'$ -эндо-конформации, то при переходе в А-форму он принимает $C3'$ -эндо-конформацию. Как было показано на рис. 11, это приводит к уменьшению расстояния между фосфатными группами и, следовательно, к уменьшению расстояний между нуклеотидными парами вдоль оси спирали (до примерно 0,25 нм при 11 нуклеотидных остатках на виток спирали).. Диаметр спирали увеличивается, изменяется ширина и глубина бороздок, пары оснований образуют с осью спирали угол около 20° и, главное, смещены к периферии спирали. Вследствие этого спи-

раль похожа на пологую винтовую лестницу, а внутри ее возникает полость диаметром 0,40 нм.

Впоследствии было обнаружено значительное число вариантов А- и В-форм ДНК (в особенности последней). Таким образом, можно говорить о двух семействах форм двойной спирали, отличающихся конформацией углеводного остатка. В то же время двойные спирали, относящиеся к любому из этих семейств, всегда закручены *вправо*.

Оказалось, однако, что наряду с правыми существуют и *левые* спирали ДНК. Первоначально левая спираль, получившая название *Z-формы* (рис. 15), была обнаружена у полинуклеотида с чередующейся последовательностью $d(GC)_n$. Цепи поли- $(dG-dC)$ самокомплементарны и в растворе с низкой ионной силой об-

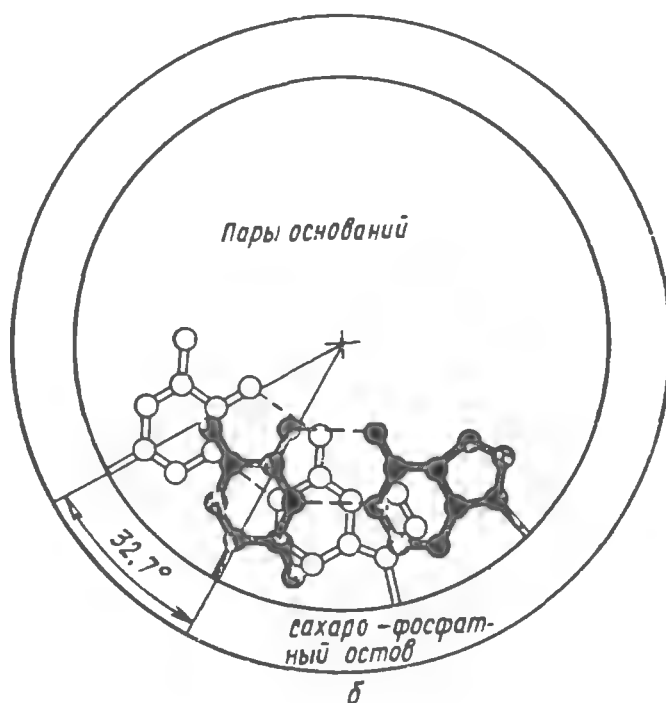
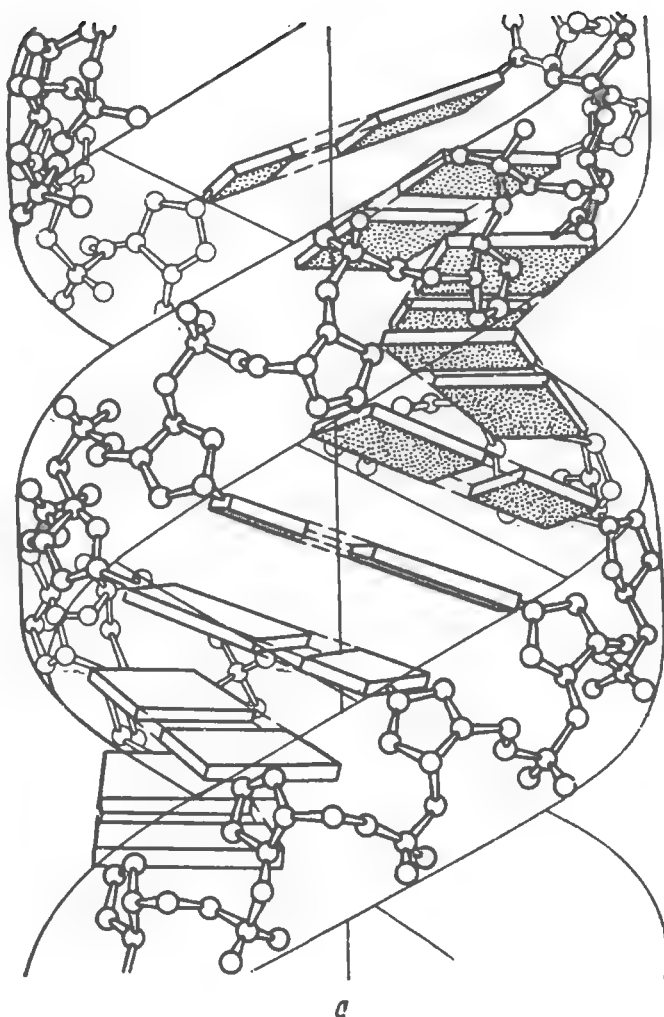


Рис. 14. А-форма двойной спирали ДНК:

а — общий вид спирали. Видно, что плоскости комплементарных пар оснований сильно наклонены к оси спирали (ср. с В-формой ДНК на рис. 8); б — вид спирали сверху. Если, мысленно двигаясь вдоль оси спирали сверху вниз, представить себе расположение последующих пар оснований, то легко заметить, что в центре двойной спирали ДНК имеется полость [На проекции, дающей вид сверху (нижний рисунок), хорошо видно, что комплементарные пары оснований расположены близко к периферии спирали.]

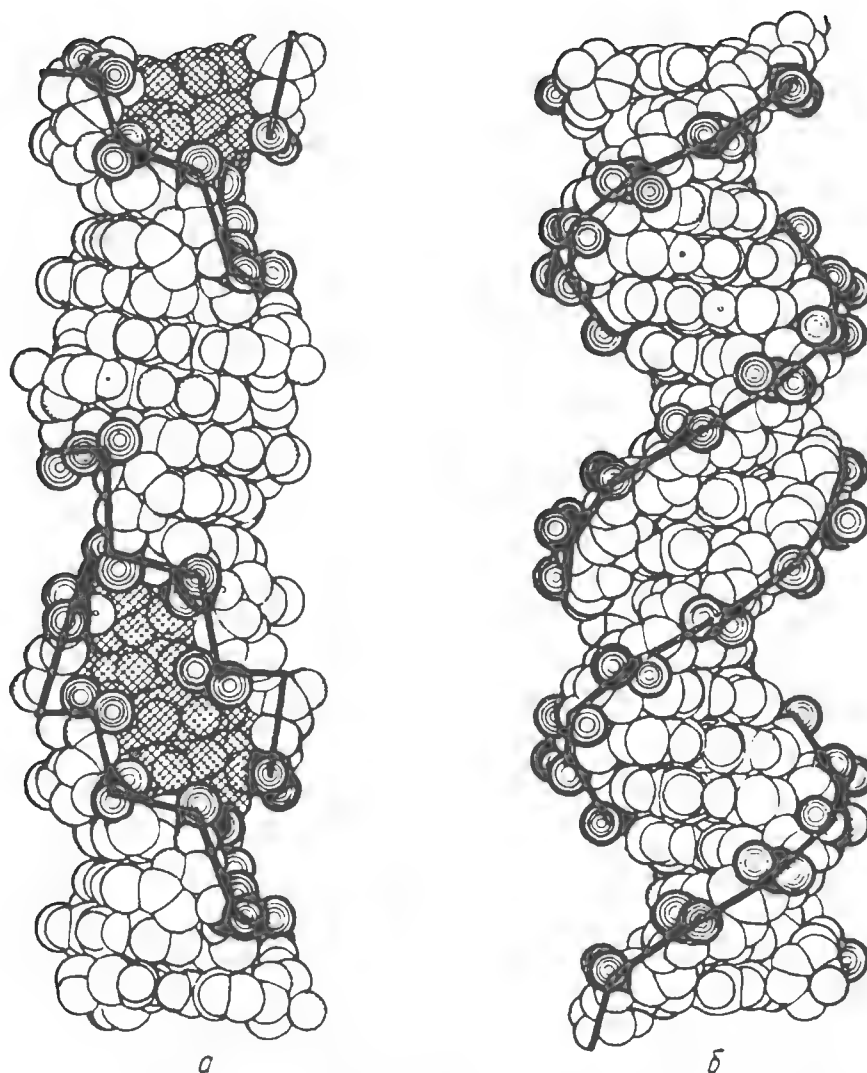


Рис. 15. Двойная спираль ДНК в Z-форме (а) и в В-форме (б)

Толстыми линиями соединены фосфатные группы (атомы которых выделены); видно, что в случае Z-формы эта линия имеет зигзагообразный вид

разуют правые двойные спирали В-типа. При высокой концентрации солей или добавлении спирта эта двойная спираль переходит в левую Z-форму. Ее замечательная особенность — чередование конформаций нуклеотидных остатков: если в dC-звеньях остатки углевода имеют C2'-эндо-, а основание — *анти*-конформации, то в dG-звеньях остатки дезоксирибозы находятся в C3'-эндо-, а основание — в *син*-конформации. В правых А- и В-формах все основания имеют *анти*-конформацию.

Таким образом, повторяющаяся единица в Z-форме ДНК состоит из двух пар нуклеотидов, а не из одной, как в В- и А-формах; в результате линия, соединяющая фосфатные группы, через каждые две пары имеет излом и принимает зигзагообразный вид (отсюда название Z-формы). По сравнению с В-формой в левой Z-форме рез-

ко изменен характер стэкинга оснований: сильные и слабые меж-плоскостные взаимодействия также чередуются.

Кроме $d(GC)_n$ -последовательности Z-форму могут принимать и другие двутяжевые полинуклеотиды или участки ДНК с чередующимися последовательностями. Важно, что в протяженной ДНК, находящейся в В-форме, принципиально возможно существование отдельных районов, имеющих Z-форму. Вопрос о существовании Z-формы ДНК *in vivo* до конца не выяснен. Однако в клетках удается обнаружить белки, которые взаимодействуют с Z-формой ДНК, но не связываются с ее В-формой.

В последние годы параметры В-, А- и Z-форм двойных спиралей ДНК удалось уточнить высокоразрешающим рентгеноструктурным анализом монокристаллов самокомплементарных синтетических олигонуклеотидов, образующих короткие двойные спирали (дуплексы). При этом можно оценить конформацию каждой нуклеотидной пары в дуплексе. Оказалось, что внутри двойной спирали существует конформационная микрогетерогенность: в зависимости от последовательности нуклеотидных пар конформации сахаров в нуклеотидных остатках несколько отличаются. Это приводит к отличиям в межнуклеотидных расстояниях вдоль оси спирали и к различному наклону пар оснований к этой оси. Такие зависящие от первичной структуры различия во вторичной структуре ДНК, по-видимому, чрезвычайно важны для ее функционирования.

г. Денатурация и ренатурация ДНК

Важнейшей особенностью внутримолекулярных взаимодействий, стабилизирующих макромолекулярную структуру ДНК, является их *кооперативность*. Для двойных спиралей, построенных из комплементарных гомополинуклеотидов, это означает, что при их денатурации, происходящей, например, при повышении температуры и приводящей к образованию изолированных полинуклеотидных цепей, все звенья спиральной структуры разрушаются одновременно. Такой процесс называют *плавлением двойной спирали*. Его описывают кривой плавления, которая, в свою очередь, характеризуется *температурой плавления* T_m и шириной температурного интервала, в котором происходит разрушение двойной спирали (рис. 16). T_m двойных спиралей, построенных только из А·Т- и только из Г·С-пар, отличаются на 40 °С. Поэтому характер кривых плавления природных ДНК сильно зависит от распределения участков, богатых А·Т- и Г·С-парами, вдоль ее молекулы. При денатурации таких ДНК, как видно из рис. 16, происходит одновременное (кооперативное) плавление отдельных протяженных сегментов ДНК. T_m ДНК сильно зависит от концентрации катионов (металлов) в растворе, экранирующих отрицательные заряды ее фосфатных групп (поэтому для большинства катионов T_m возрастает с ростом их концентрации).

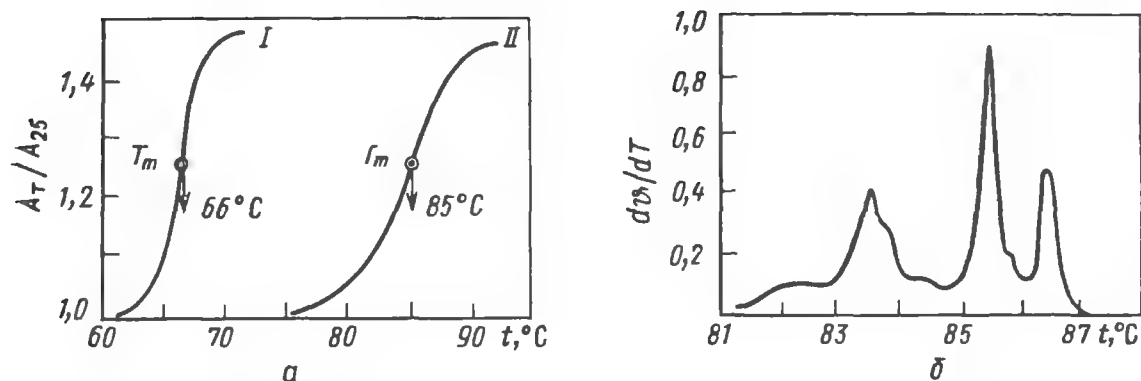


Рис. 16. Плавление ДНК:

а — кривые плавления двуспирального комплекса поли(дТ) — поли(дА) (I) и ДНК бактерии *Proteus vulgaris* (II). [Денатурация ДНК сопровождается увеличением их оптической плотности в УФ-области (около 260 нм), называемым гиперхромным эффектом. Это позволяет следить за ходом денатурации ДНК.]: A_T/A_{25} — отношение оптической плотности ДНК при данной температуре к оптической плотности нативной ДНК (25 °C); б — дифференциальная кривая плавления фрагмента ДНК бактериофага fd. [Ее получают путем тщательного измерения зависимости Φ (доли звеньев ДНК, перешедших в денатурированное состояние) от температуры.]: каждый пик на кривой соответствует выплавлению отдельного участка ДНК

В молекулярной биологии широко используется способность денатурированных ДНК ренатурировать с восстановлением исходной двуспиральной структуры. Она лежит в основе метода *молекулярной гибридизации* нуклеиновых кислот, который позволяет выявлять степень сходства различных ДНК (а также РНК). Для этого денатурированную ДНК (если изучается гибридизация двух различных нуклеиновых кислот, то одна из них несет радиоактивную метку) помещают в условия, оптимальные для образования двойных спиралей (ионная сила раствора — около 0,2; температура — на 10—20 °C ниже T_m нативной ДНК). В случае полностью комплементарных цепей ДНК со временем они целиком превратятся в двуспиральные молекулы. Если в смеси присутствуют как комплементарные, так и некомплементарные цепи ДНК, то после ренатурации первых тем или иным способом определяют долю двуспиральных молекул. В настоящее время широко распространены методы, когда денатурированные молекулы ДНК одного типа закрепляются на нитроцеллюлозных фильтрах, которые затем помещают в раствор ДНК (или РНК) другого типа. После образования двуспиральных комплексов на фильтрах они легко могут быть отмыты от несвязавшейся ДНК. Этот же подход используется при выявлении цепей ДНК (или РНК), комплементарных другим ДНК (или РНК), после разделения их электрофорезом в гелях.

Скорость процесса ренатурации ДНК обычно описывают уравнением

$$\frac{C}{C_0} = \frac{1}{1 + kC_0t},$$

где C — концентрация однотяжевой ДНК; C_0 — суммарная концентрация ДНК; t — время; k — константа скорости второго по-

рядка. Из этого уравнения видно, что доля денатурированной ДНК, еще не реассоциированной к моменту времени t , есть функция $C_0 t$. Поэтому процесс ренатурации ДНК обычно описывают в координатах, показанных на рис. 17, а его важнейшей характеристикой является величина $C_0 t_{1/2}$, при которой реассоциация проходит на 50 %.

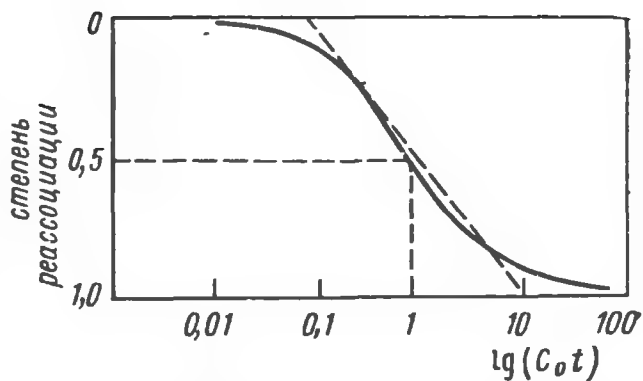


Рис. 17. Идеализированная кривая реассоциации цепей денатурированной ДНК (см. также рис. 107 и пояснения к нему в тексте)

д. Сверхспирализация ДНК; топоизомеразы

В середине 1960-х годов Д. Виноград и сотр. обнаружили, что ДНК некоторых бактериофагов и митохондрий может существовать в виде циклических молекул. Позже было установлено, что большинство вирусных и множество клеточных ДНК имеют кольцевую форму. В том случае, если обе полинуклеотидные цепи в кольцевой молекуле, образованной двуспиральной ДНК, ковалентно замкнуты (аналогичная ситуация возникает, когда концы петель двуспиральной ДНК скреплены белками), то они уже не могут быть разделены в пространстве.

Важнейший параметр каждой кольцевой ковалентно замкнутой молекулы ДНК — порядок зацепления двух одноцепочечных колец в ней. Он обозначается как Lk (или α) и в первом приближении равен числу пересечений одной полинуклеотидной цепи с другой в двуспиральном кольце (рис. 18). Lk , таким образом, имеет целочисленные значения. Самое важное, Lk — величина постоянная (инвариантная) для данной ковалентно замкнутой кольцевой ДНК.

Если такая ДНК находится в так называемом *релаксированном состоянии*, т. е. без всякого напряжения может быть расположена в одной плоскости, то $Lk = Tw$, где Tw (обозначаемое также β^0) равно числу витков в двойной спирали данной ДНК.

Однако если увеличить или уменьшить число витков двойной спирали, сохранив ковалентную непрерывность обеих цепей, то в силу инвариантности Lk такое изменение будет компенсироваться образованием *сверхвитков* (рис. 18). Число сверхвитков в сверхспирализованной ДНК обозначается Wr (или τ ; его часто называют «райзингом» от англ. writhe — скрючиваться) и определяется как

$$Wr = Lk - Tw.$$

Wr может быть как положительной, так и отрицательной величиной. Кольцевые ковалентно замкнутые ДНК, выделенные из

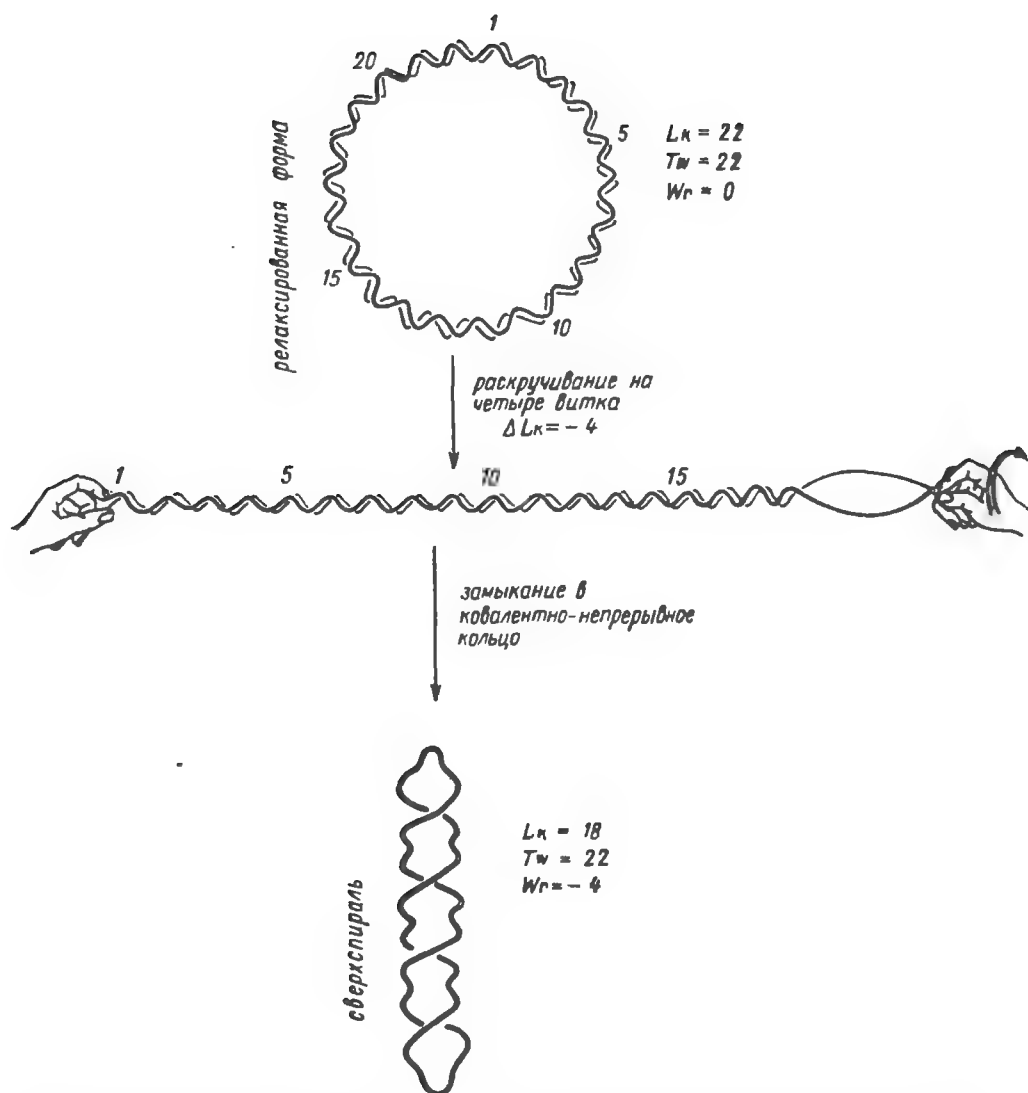


Рис. 18. Схема, иллюстрирующая образование отрицательно сверхспирализованной кольцевой ковалентной замкнутой ДНК

клеток, всегда имеют отрицательную сверхспирализацию. Степень сверхспирализации той или иной ДНК удобно характеризовать плотностью сверхвитков σ , которая равна Wr/Tw . Для многих природных сверхспирализованных ДНК σ равно примерно $-0,05$.

С функциональной точки зрения важно, что сверхспирализованная ДНК обладает значительным запасом энергии по сравнению с ее релаксированной формой. Следовательно, локальное расплетание двойной спирали ДНК с отрицательными сверхвитками будет приводить к сбросу напряжения сверхспирализации, и потому оно энергетически выгодно. Это отчетливо проявляется в том, что отрицательная сверхспирализация заметно стимулирует переход ДНК из правой В-формы в левую Z-форму. Действительно, уже при обычных физиологических условиях участки с последовательностями $d(CG)_n \cdot d(CG)_n$ и $d(AC)_n \cdot d(GT)_n$, встроенными в ДНК с $\sigma = -0,06$, переходят в левоспиральную Z-форму. Есть веские свидетельства

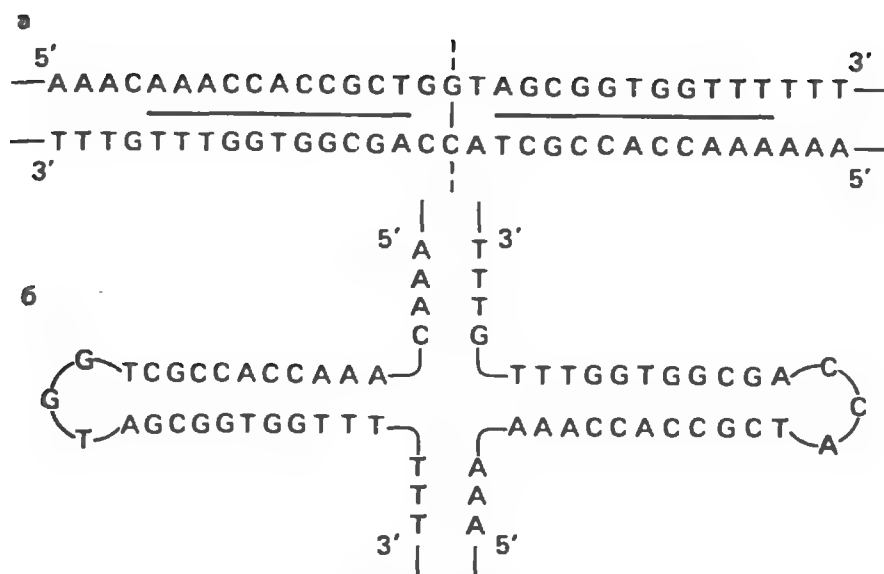


Рис. 19. Линейная (а) и крестообразная (б) структуры участка плазмиды pBR322, содержащего обращенную повторяющуюся последовательность нуклеотидных пар, подчеркнутую на рисунке

Пунктиром показано положение пары, через которую проходит ось симметрии второго порядка; поворот вокруг этой оси на 180° в плоскости чертежа приводит к совмещению повторяющихся последовательностей

тому, что отрицательная сверхспирализация ДНК способствует существованию некоторых ее участков в Z-форме *in vivo*.

Другой пример влияния сверхспирализации на структурные превращения двойной спирали ДНК — образование крестообразных структур. Практически любая ДНК содержит инвертированные, или палиндромные, повторяющиеся последовательности длиной от нескольких п. о. до многих тысяч п. о. Теоретически можно представить себе превращение линейной двуспиральной формы палиндрома в крестообразную (рис. 19). Для релаксированной ДНК вероятность такого превращения ничтожна. Поскольку в ДНК с отрицательными сверхвитками этот переход энергетически выгоден, крестообразные формы *in vitro* обнаруживаются у всех исследованных сверхспирализованных ДНК с нормальной плотностью сверхвитков. (Экспериментально крестообразные структуры фиксируют по наличию однотяжевых петель в вершине «шпилек», которые расщепляются нуклеазами, специфичными к однотяжевой ДНК.) Вопрос о существовании крестообразных структур ДНК *in vivo* остается открытым. Скорость их образования очень мала, и, может быть, именно поэтому в клетке их еще никому не удалось обнаружить.

Если ДНК содержит гомопиримидин — гомопуриновые последовательности, то под влиянием отрицательной сверхспирализации она может переходить в форму, показанную на рис. 20. Поскольку для последовательностей типа $d(AG)_n \cdot d(CT)_n$ такой переход происходил при пониженных значениях pH (а при pH 4,3 он наблюдается и при $\sigma=0$), она получила название H-формы. Наличием H-фор-

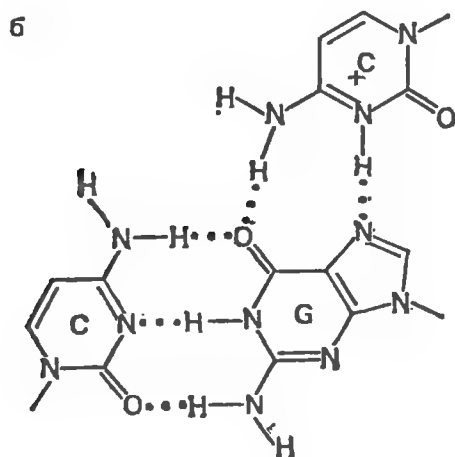
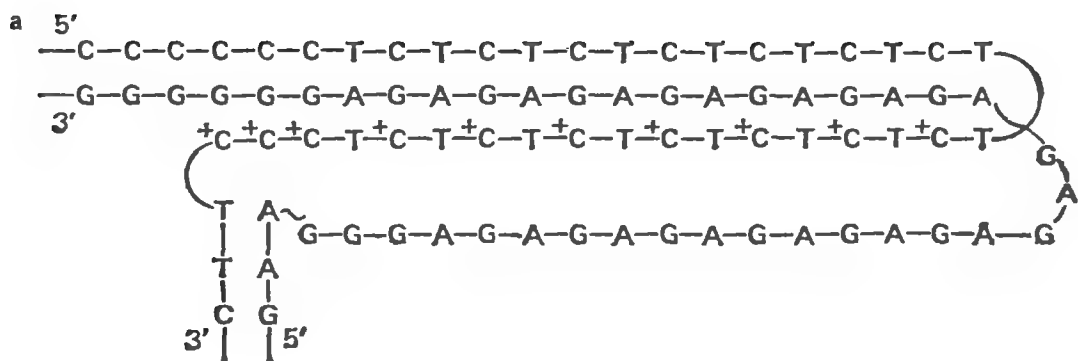


Рис. 20. Н-форма ДНК, возникающая при встраивании показанной на рисунке последовательности в отрицательно сверхспирализованную ДНК (при рН 4,3) (а) ; строение триплета оснований в тройной спирали Н-формы ДНК (б)

(Обратите внимание на то, что один из остатков С протонирован.)

мы объясняют существование в природных ДНК областей, сверхчувствительных к нуклеазам, специфичным к однотяжевым полинуклеотидам.

Отметим также, что отрицательная сверхспирализация ДНК облегчает связывание с ней белков, раскручивающих ее двойную спираль.

В клетке отрицательная сверхспирализация ДНК создается ДНК-гиразами, относящимися к классу топоизомераз. *Топоизомеразы* — это ферменты, способные изменять число сверхвитков в кольцевых ковалентно замкнутых ДНК. Они подразделяются на две группы — топоизомеразы I и II. Субстрат для топоизомеразы I — сверхспирализованная ДНК, в которой они уменьшают число сверхвитков, превращая ее в конечном итоге в релаксированную кольцевую ДНК. Топоизомеразы II (ДНК-гиразы), напротив, способны создавать сверхвитки ДНК. При этом каждый акт работы ДНК-гиразы сопровождается созданием двух отрицательных сверхвитков, а ее активность сопряжена с гидролизом АТР.

Как это будет ясно из последующих глав, топоизомеразы играют исключительно важную роль в процессах репликации и рекомбинации ДНК. Они содержатся в каждой про- и эукариотической клетке.

Макромолекулярная структура РНК

а. Двухтяжевые РНК

РНК, построенные из двух комплементарных полинуклеотидных цепей, встречаются редко и обнаруживаются только в составе некоторых вирусов. Кроме того, известны однотяжевые РНК (вирионы, например; см. гл. 13), состоящие, по сути дела, из двух комплементарных нуклеотидных последовательностей и поэтому существующие в виде двуспиральных молекул.

В двойной спирали РНК возможна только одна — $C3'$ -эндо-конформация остатка рибозы ($C2'$ -эндо-конформация запрещена из-за слишком тесных контактов между $2'$ -гидроксилами и другими атомами). Поэтому двуспиральные РНК, а также комплементарные двухтяжевые полирибонуклеотидные комплексы [типа поли(А).

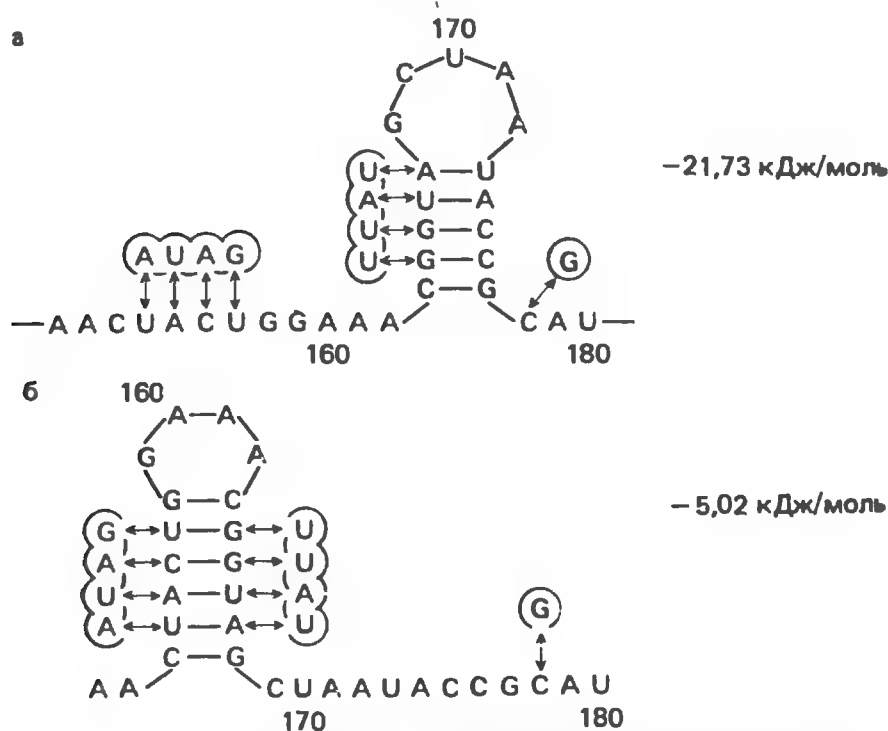
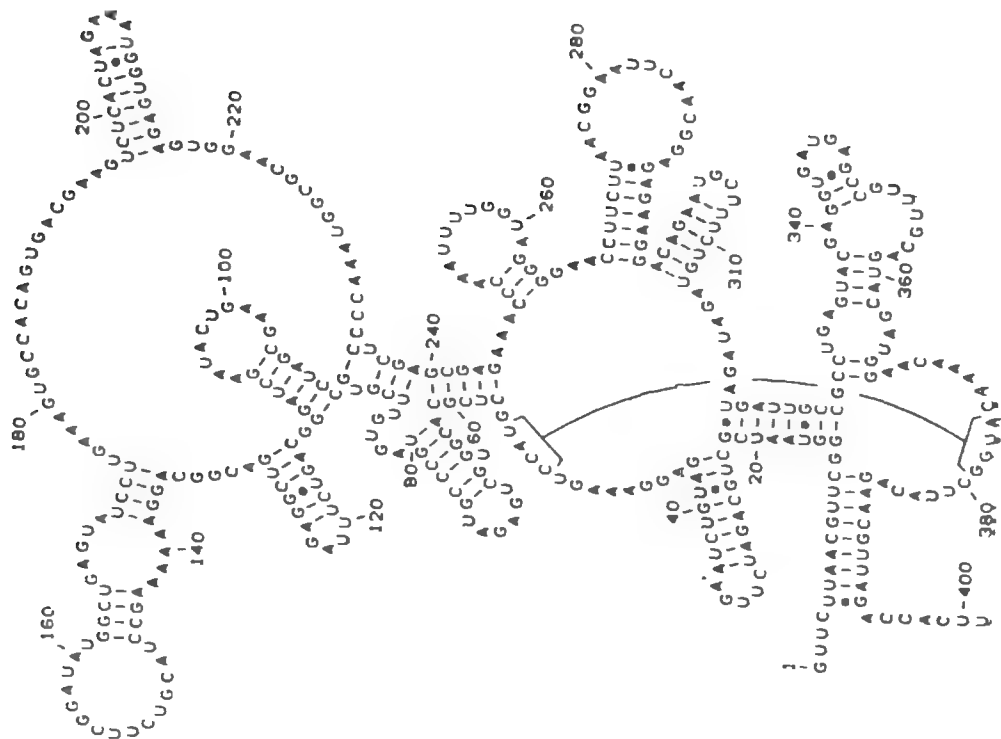
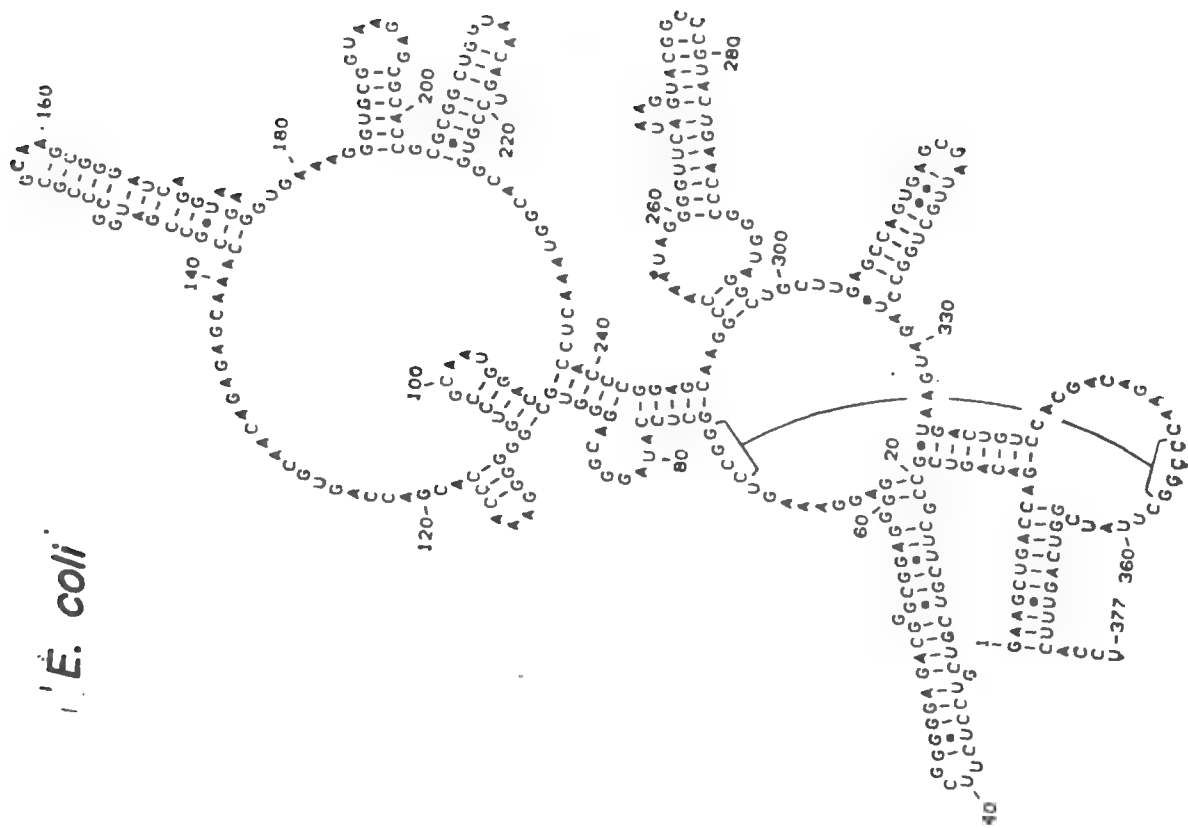


Рис. 21. Схема, иллюстрирующая применение сравнительного («филогенетического») анализа для выбора модели вторичной структуры РНК

Показаны два варианта возможной вторичной структуры участка рибосомной 16S РНК *E. coli* (номера-расстояния в нуклеотидных остатках от 5'-конца РНК). Несмотря на то что энергия образования структуры а ниже, предпочтение отдается структуре б, поскольку замены в нуклеотидной последовательности аналогичного участка в молекуле рибосомной 16S РНК *Bac. brevis* (показаны стрелками) не допускают существования структуры а. В то же время в случае структуры б, несмотря на нуклеотидные замены, двухтяжевой участок сохраняется в обеих РНК (такие замены называются компенсаторными).

Bac. subtilis

E. coli



•поли(U)] существуют только в правой А-форме, в основных чертах подобной А-форме ДНК. В то же время принципиально возможно существование и левых двойных спиралей РНК (Z-РНК).

б. Общие черты вторичной структуры однотяжевых РНК

Макромолекулы большинства природных РНК построены из одной полирибонуклеотидной цепи. Основным элементом их вторичной структуры — сравнительно короткие двойные спирали, образованные комплементарными участками одной и той же цепи и перемежающиеся ее однотяжевыми сегментами. Полирибонуклеотидные цепи в таких двуспиральных структурах антипараллельны, а сами двойные спирали, находящиеся в А-форме, не идеальны: в них имеются дефекты в виде неспаренных нуклеотидных остатков или не вписывающихся в двойную спираль однотяжевых петель (рис. 21 и 22). Наряду с классическими уотсон-криковскими парами (А·U и G·C) в двутяжевых участках РНК часто встречается пара G·U. Таким образом, стабильность двутяжевых районов поддерживается комплементарными и межплоскостными взаимодействиями оснований. В однотяжевых участках наблюдаются сильные стэкинг-взаимодействия оснований, вследствие чего они стремятся принять конформацию однотяжевой спирали.

Общие принципы организации вторичной структуры РНК были сформулированы П. Доти и сотр. и А. С. Спириным и сотр. (1959—1961). Впоследствии было установлено, что эти принципы справедливы для всех однотяжевых РНК и на их основе для любой РНК с известной первичной структурой может быть построена достаточно достоверная модель вторичной структуры.

Поиск такой модели начинают с выявления с помощью ЭВМ вариантов вторичных структур с числом комплементарных пар оснований, близким к максимальному. Далее отбираются вторичные структуры, образование которых соответствует минимуму свободной энергии. Для этого используют специальные алгоритмы, которые позволяют учесть протяженность двуспиральных участков, количество и последовательность чередования G·C-, A·U- и G·U-пар в них, характер и размеры дефектов в этих участках и т. д. Такие алгоритмы создают на базе экспериментальных данных, полученных при изучении стабильности большого числа синтетических олигонуклеотидных комплексов, моделирующих одно- и двутяжевые участки РНК.

Рис. 22. Вторичные структуры РНК-компонента рибонуклеазы Р из *E. coli* и *Bac. subtilis*

Несмотря на многочисленные нуклеотидные замены, каркас структуры сохраняется. Показаны также участки молекул (одинаково расположенные), способные образовать «псевдоузел» (см. рис. 25). Обе РНК обладают ферментативной активностью и выполняют одну и ту же функцию — участвуют в процессинге тРНК

К сожалению, подобные расчеты дают для данной РНК, как правило, несколько (а для высокомолекулярных РНК — множество) моделей вторичных структур с близким значением свободной энергии. Однако теоретический поиск модели может быть продолжен, если известны нуклеотидные последовательности хотя бы еще одной (лучше нескольких) РНК, относящейся к тому же классу, что и исследуемая, но выделенной из другого объекта. В этом случае с успехом применяют так называемый «филогенетический» подход, который основан на допущении, что РНК, выполняющие в клетке одну и ту же функцию, должны иметь сходные вторичные и третичные структуры. Тогда среди расчетных моделей для двух (или большего числа) РНК нужно выбрать однотипные.

Впервые правомерность «филогенетического» подхода была продемонстрирована при изучении структуры тРНК. Оказалось, что вторичная структура любых тРНК независимо от вариаций в их нуклеотидных последовательностях описывается универсальной моделью «клеверного листа». Еще более впечатляют результаты применения этого подхода к рибосомным РНК, когда из многих сотен расчетных моделей удалось выбрать по одной для РНК каждого типа *. Как осуществляется такой выбор на уровне отдельных элементов вторичных структур РНК, показано на рис. 21, а на рис. 22 демонстрируется подобие вторичных структур РНК-компонента рибонуклеазы Р (о его ферментативной активности см. в гл. VIII).

Теоретическая модель вторичной структуры РНК должна быть далее подвергнута экспериментальной проверке. Прямые методы определения конформации макромолекул — рентгеноструктурный анализ и ядерный магнитный резонанс (ЯМР) — пока применимы лишь для низкомолекулярных РНК (см. следующий раздел). Поэтому в большинстве случаев принадлежность того или иного нуклеотидного остатка РНК (или достаточно протяженных ее участков) к двуспиральному или однотяжевому элементу вторичной структуры оценивается косвенным путем. Основная роль здесь принадлежит методам химической модификации и методам расщепления РНК структуроспецифическими РНКазами.

Действительно, для каждого гетероциклического основания можно подобрать такой химический реагент, который избирательно взаимодействует только с атомами или группами, участвующими в образовании водородных связей при комплементарном спаривании нуклеотидных остатков. Так, например, кетоксаль избирательно взаимодействует с N1 и 2-NH₂-группой гуанина, диметилсульфат (в определенных условиях) — с N1 аденина и N3 цитозина, карбодимид — с N3 урацила. Следовательно, если тот или иной нуклео-

* Модели вторичных структур транспортных и рибосомных РНК подробно рассмотрены во втором томе этого учебника (Спирин А. С. Молекулярная биология: Структура рибосом и биосинтез белка. — М.: Высшая школа, 1986).

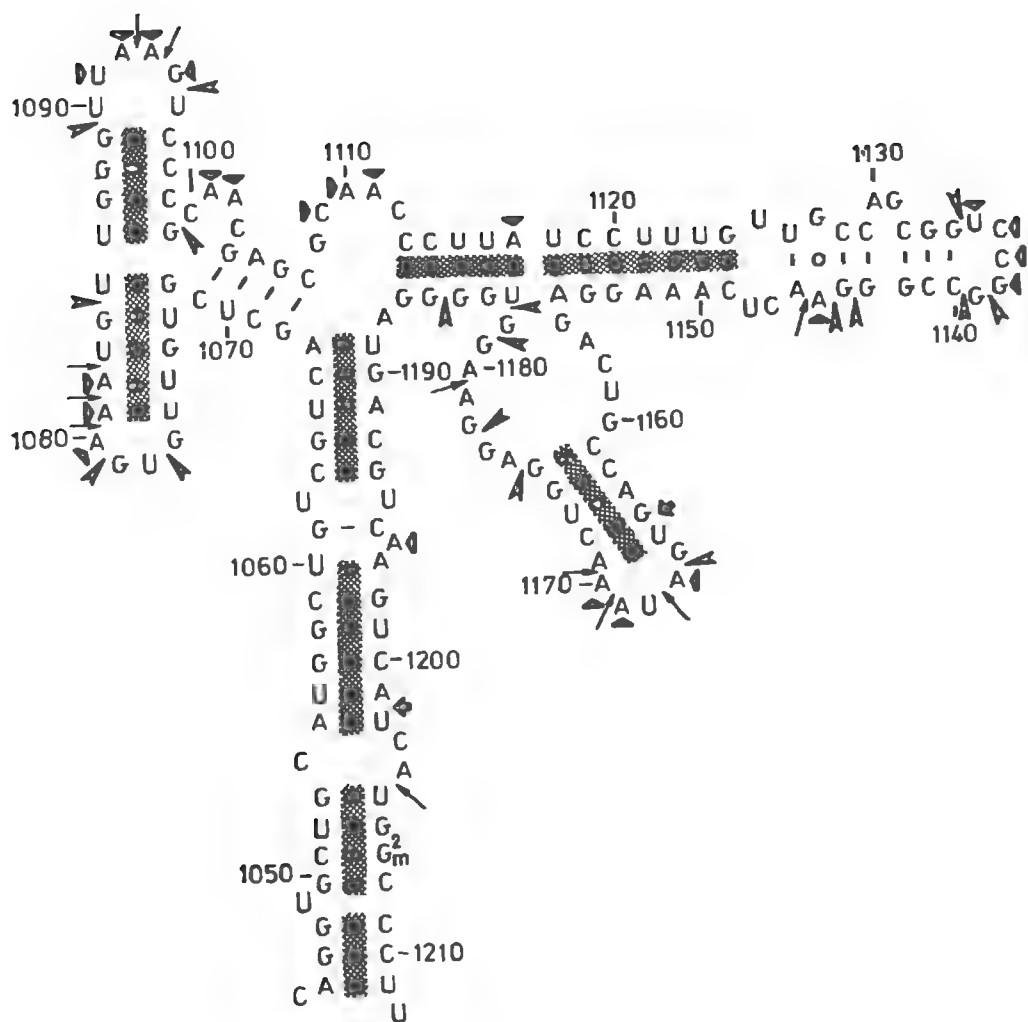


Рис. 23. Схема, иллюстрирующая экспериментальную проверку модели вторичной структуры РНК

Приведен фрагмент рибосомной 16 S РНК. Выделены (штриховкой) двуспиральные участки, существование которых подтверждается сравнительным анализом (см. рис. 21). Стрелками показаны места расщепления РНК гуаниловой T_1 (зачерненная галка), адениловой U_2 (тонкая стрелка) РНКазами и нуклеазой из яда кобры (жирная стрелка), а также нуклеотидные остатки, модифицируемые химическими реагентами, специфичными к однотяжевым участкам РНК (зачерненный треугольник). Эти результаты в основном подтверждают модель, однако существование двуспирального участка, примыкающего к A1080, маловероятно, так как почти вся его «левая» цепь модифицируется (или расщепляется) агентами, специфичными к однотяжевой РНК

тидный остаток РНК модифицируется одним из этих реагентов, он принадлежит однотяжевому участку РНК. Отсутствие модификации интерпретировать труднее — нуклеотидный остаток может входить в состав двойной спирали, но может быть экранирован и третичными взаимодействиями (см. следующий раздел).

Аналогичные принципы лежат в основе ферментативных методов анализа вторичных структур РНК. Известно большое число РНКаз, которые направленно гидролизуют фосфодиэфирные связи в однотяжевых участках РНК. Обнаружено также несколько нуклеаз,

имеющих повышенное сродство к двутяжевым РНК. Анализируя картину гидролиза РНК этими ферментами, можно установить принадлежность определенных нуклеотидных остатков к тому или иному типу вторичной структуры. Результаты, получаемые химическими и ферментативными методами, хорошо дополняют друг друга, что видно из рис. 23.

Идентификация модифицированных нуклеотидных остатков в полинуклеотидной цепи РНК долгое время была задачей особой трудности. С появлением современных методов секвенирования нуклеиновых кислот она существенно упростилась. Модификацию РНК или ее расщепление ферментами ведут таким образом, чтобы (как и при секвенировании) было затронуто в среднем только одно звено на молекулу (в чем есть дополнительный смысл, так как множественная модификация РНК искажает ее структуру). Далее, если изучается РНК небольшого размера или сегмент РНК, примыкающий к одному из ее концов, то этот конец метят радиоактивной меткой и задача идентификации модифицированного основания (после расщепления соответствующего звена) или атакованной нуклеазой межинуклеотидной связи сводится, как и при секвенировании, к определению длины фрагмента по его подвижности в высокоразрешающем электрофорезе в геле. В том случае, когда анализируемый район удален от концов молекулы на расстояние больше 150—200 н. о., используют реакцию обратной транскрипции (см. гл. 13). Для этого синтезируют олигонуклеотид, комплементарный участку РНК, расположенному вблизи от анализируемого района с 3'-концевой стороны молекулы, и далее используют его как праймер для обратной транскриптазы. Так как этот фермент «останавливается» на модифицированных остатках матрицы (или в том месте, где расщеплена фосфодиэфирная связь), то вновь по длине образующегося фрагмента можно определить положение модифицированного звена в РНК.

С помощью этих подходов в настоящее время проанализированы модели вторичных структур десятков вирусных и информационных РНК.

в. Третичная структура однотяжевых РНК

Уже в ранних исследованиях макромолекулярной организации однотяжевых РНК было установлено, что в физиологических условиях они характеризуются компактной и упорядоченной третичной структурой, которая возникает за счет взаимодействия шпилькообразных элементов их вторичной структуры. Транспортные РНК — единственные представители природных полирибонуклеотидов, которые удалось закристаллизовать и изучить методом рентгеноструктурного анализа с достаточно высоким разрешением. Поэтому представления о принципах организации третичной структуры РНК

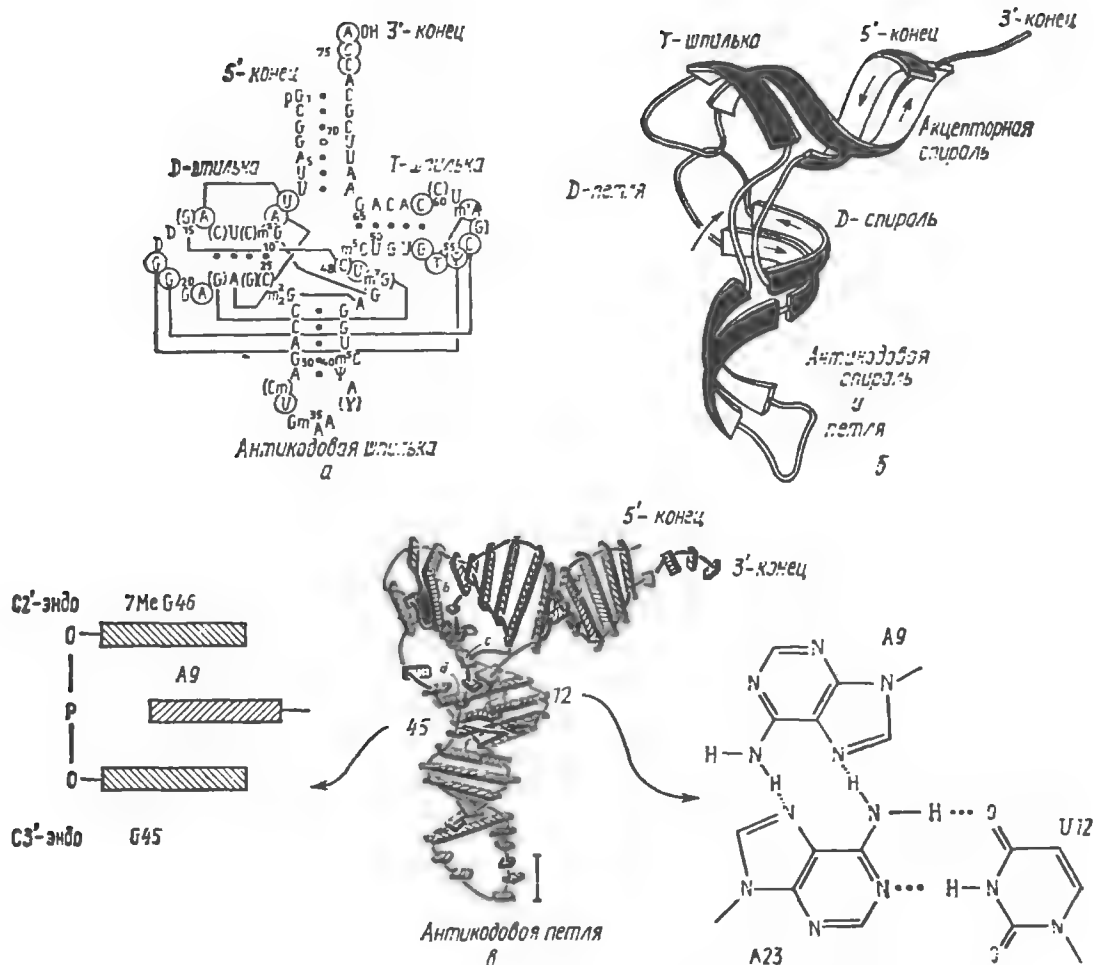


Рис. 24. Вторичная и третичная структуры фенилаланиновой транспортной РНК дрожжей:

а — «клеверный лист»; тонкие линии соединяют основания, которые образуют «третичные» пары в триплетах в макромолекуле тРНК; *б* — модель пространственной структуры тРНК; лентами показаны двуспиральные участки РНК; *в* — модель третичной структуры тРНК, демонстрирующая расположение плоскостей оснований в нуклеотидных парах и триплетах, а также в одиотяжевых участках; слева — схема интеркалляции основания одного участка цепи между двумя основаниями другого; справа — пример «третичного» контакта другого типа — А·U·А-триплета

практически полностью основаны на данных о пространственной структуре тРНК (рис. 24).

Коротко рассмотрим эти принципы.

1. Элементы вторичной структуры РНК располагаются друг относительно друга так, чтобы обеспечить максимальный стэкинг оснований в макромолекуле в целом.

2. Контакты между отдельными элементами вторичной структуры осуществляются за счет нескольких типов так называемых «третичных» внутримолекулярных взаимодействий:

а) за счет образования дополнительных, зачастую неутсон-криковских пар оснований между нуклеотидными остатками удаленных друг от друга (в первичной и вторичной структурах) одиотяжевых участков и триплетов оснований между нуклеотидными ос-

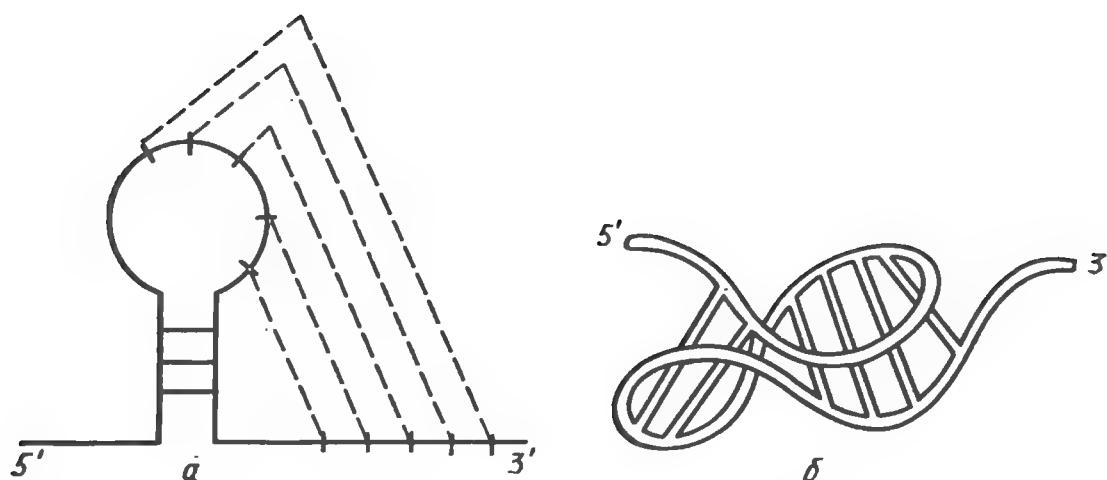


Рис. 25. Схема образования «псевдоузла» благодаря спариванию комплементарных участков РНК в вершине шпильки и в другом однотяжевом участке (а); пространственная модель «псевдоузла» (б)

татками однотяжевых и двутяжевых элементов; в последнем случае в образовании водородных связей участвуют N7-атомы пуриновых оснований, а также функциональные группы при их C6-атомах, уже занятые в образовании одной водородной связи в уотсон-криковской паре. При анализе вторичных структур вирусных и рибосомных РНК было обнаружено, что однотяжевые участки в вершине шпилек в принципе могут давать двуспиральные комплексы с межшпилечными однотяжевыми сегментами; если это действительно так, то в молекулах РНК должны существовать так называемые «псевдоузлы», схематически показанные на рис. 25;

б) за счет дополнительных («третичных») стэкинг-взаимодействий после интеркаляции (внедрения) оснований одного участка между двумя соседними основаниями другого однотяжевого участка (рис. 24);

в) за счет образования дополнительных водородных связей между 2'-ОН-группами остатков рибозы и основаниями, а также другими группами сахарофосфатного остова.

3. Третичная структура РНК стабилизирована ионами двухвалентных металлов, которые связываются не только с фосфатными группами, но и с основаниями.

Нет сомнений в том, что аналогичные принципы лежат в основе организации структуры всех однотяжевых РНК. Однако строить конкретные пространственные модели макромолекул однотяжевых РНК крайне затруднительно из-за множества возможных вариантов укладки в них элементов вторичной структуры. Некоторую дополнительную информацию дает здесь идентификация нуклеотидных остатков, которые могут быть сшиты бифункциональными химическими агентами.

ЛИТЕРАТУРА

Шабарова З. А., Богданов А. А. Химия нуклеиновых кислот и их компонентов.— М.: Химия, 1978.

Зенгер В. Принципы структурной организации нуклеиновых кислот/Пер. с англ.—М.: Мир, 1987.

Кантор Ч., Шиммель П. Биофизическая химия/Пер. с англ.— М.: Мир, 1984. Т. 1.

Спирин А. С. Молекулярная биология: Структура рибосомы и биосинтез белка.— М.: Высшая школа, 1986.

Уотсон Дж., Туз Дж., Курц Д. Рекомбинантные ДНК/Пер. с англ.— М.: Мир, 1986.

Льюин Б. Гены/Пер. с англ.—М.: Мир, 1987.

Watson J. D., Crick F. H. C. A structure of deoxyribose nucleic acid//Nature, 1953. V. 171. P. 737—738.

Часть

Б

Биосинтез нуклеиновых кислот

Глава II

РЕПЛИКАЦИЯ ДНК

Способность генетического материала, ДНК, к самовоспроизведению (репликации) лежит в основе размножения живых организмов, передачи наследственных свойств из поколения в поколение и развития многоклеточного организма из зиготы. Настоящая глава посвящена молекулярным механизмам самовоспроизведения ДНК.

1. МАТРИЧНЫЙ СИНТЕЗ

Модель ДНК Уотсона и Крика сразу же позволила понять принцип удвоения ДНК. Поскольку каждая из цепей ДНК содержит последовательность нуклеотидов, комплементарную другой цепи, т. е. их информационное содержание идентично, представлялось вполне логичным, что при удвоении ДНК цепи расходятся, а затем каждая цепь служит матрицей, на которой выстраивается комплементарная ей новая цепь ДНК. В результате образуются два дуплекса ДНК, каждый из которых состоит из одной цепи исходной родительской молекулы ДНК и одной новосинтезированной цепи. Экспериментально показано, что именно так, по *полуконсервативному механизму*, происходит репликация ДНК (рис. 26). Несмотря на простоту основного принципа, процесс репликации сложно организован и требует участия множества белков. Эти белки, как и все другие, закодированы в последовательности нуклеотидов ДНК. Таким образом, возникает важнейшая для жизни петля обратной связи: ДНК направляет синтез белков, которые реплицируют ДНК.

ДНК-полимеразы

Комплементарное копирование матрицы осуществляют ферменты ДНК-зависимые ДНК-полимеразы или просто *ДНК-полимеразы*. Первый фермент этого типа был открыт в 1956 г. у *E. coli*, затем ДНК-полимеразы были обнаружены в других организмах. ДНК-полимеразы ведут синтез ДНК на одноцепочечной матрице, если имеется затравка — комплементарный матрице фрагмент растущей

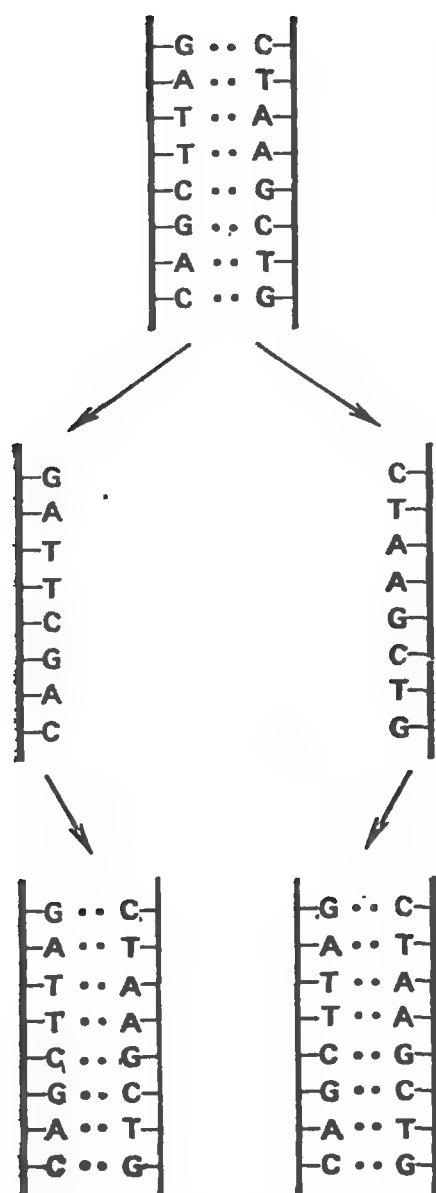


Рис. 26. Полуконсервативная репликация

В ходе репликации две комплементарные цепи ДНК разделяются и каждая из них служит матрицей для синтеза новой комплементарной цепи (тонкая линия). В результате образуются одинаковые дочерние молекулы, идентичные исходной молекуле ДНК

цию в активированной высокоэнергетической форме дезоксирибонуклеозидтрифосфата. В этом отношении синтез ДНК напоминает синтез всех других биополимеров: поскольку полимеризация мономеров в полимер энергетически не выгодна, мономеры всегда поступают в реакцию синтеза в активированной форме. В случае синтеза ДНК присоединение очередного нуклеотида к концу затравки сопровождается гидролизом богатой энергией связи и отщеп-

цепи ДНК. ДНК-полимеразы последовательно наращивают конец затравки, шаг за шагом присоединяя к нему следующие нуклеотиды (рис. 27), причем выбор очередного нуклеотида для присоединения к концу затравки диктуется матрицей.

Очередной нуклеотид, субстрат для ДНК-полимеразы, поступает в реак-

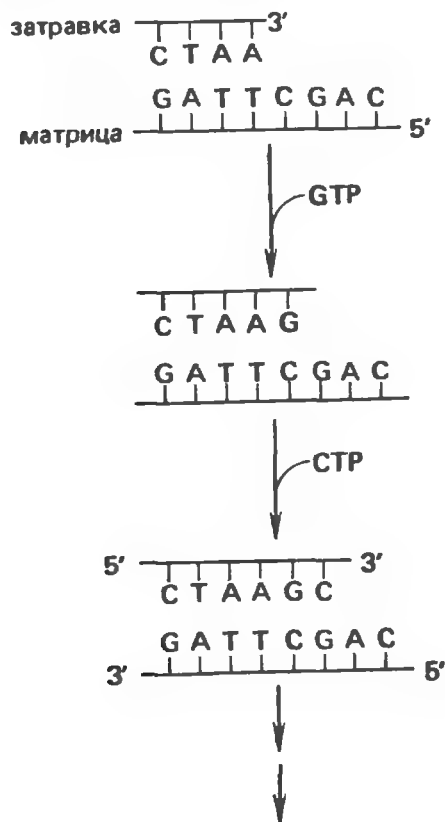


Рис. 27. Матричный синтез ДНК

ДНК-полимераза шаг за шагом удлиняет затравку. Цепи матрицы и затравки направлены в противоположные стороны — антипараллельны

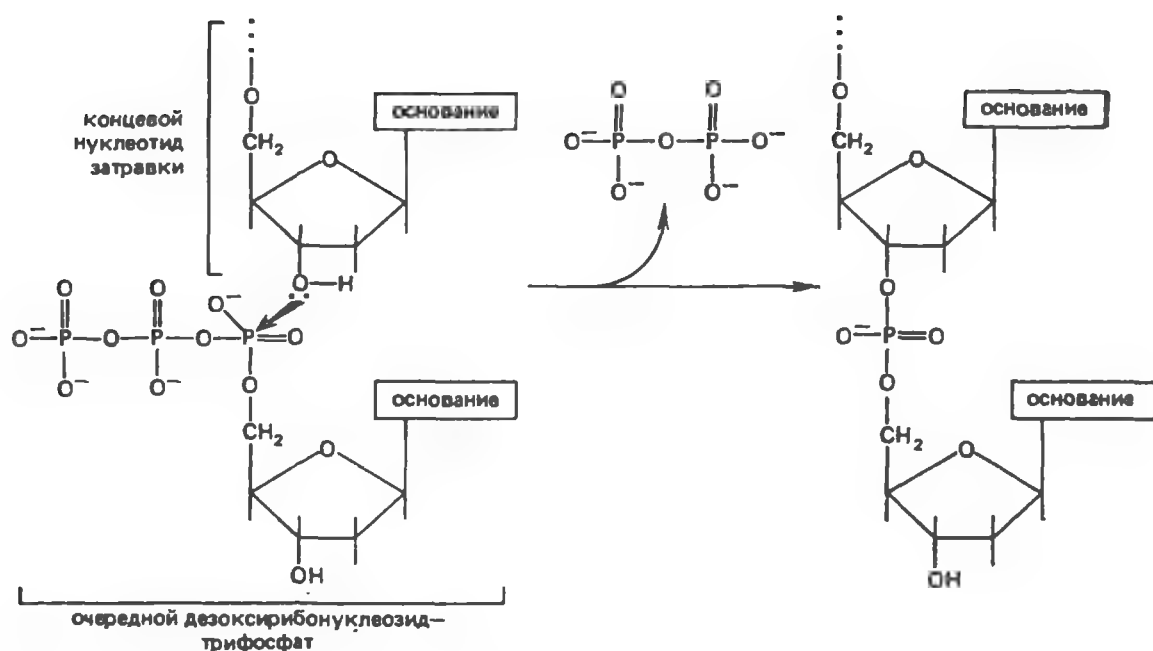


Рис. 28. Реакция, осуществляемая ДНК-полимеразой

Фермент катализирует нуклеофильную атаку 3'-гидроксильной группы затравки, направленную на α-фосфат очередного дезоксирибонуклеозидтрифосфата (в том случае, если он комплементарен матрице; матрица на рисунке не показана)

лением пирофосфата, что и делает реакцию в целом энергетически выгодной. Наличие в клетке пирофосфатазы обеспечивает расщепление пирофосфата и делает реакцию практически необратимой. При полимеризации растет всегда 3'-конец затравки, т. е. синтез происходит в направлении 5'→3': 3'-ОН-группа концевой нуклеотида затравки атакует α-фосфат очередного дезоксирибонуклеозидтрифосфата (но только в том случае, если он комплементарен очередному нуклеотиду матрицы!), в результате чего отщепляется пирофосфат, а дезоксирибонуклеозидмонофосфат оказывается связанным фосфодиэфирной связью с растущей цепью ДНК, удлиняя ее на одно звено (рис. 28).

Затравка антипараллельна матрице (рис. 27). Естественно, эта полярность сохраняется и при ее дальнейшем росте, так что результатом работы ДНК-полимеразы на одноцепочечной матрице является антипараллельная двойная спираль ДНК.

Необходимая для полимеризации геометрия химических группировок в активном центре ДНК-полимераз задается с помощью координационного взаимодействия определенных атомов матрицы, затравки и субстрата с ионами металлов. Субстраты поступают в реакцию синтеза ДНК в виде комплексов (хелатов) с ионами Mg^{2+} . Ионы Zn^{2+} , входящие в активный центр почти всех матричных ферментов, как ДНК-, так и РНК-полимераз, обеспечивают правильную взаимную ориентацию 3'-ОН-группы затравки, α-фосфата субстрата и очередного нуклеотида матрицы. Возможно, это обстоятельство отражает историю возникновения первых систем мат-

ричного синтеза полинуклеотидов; показано, что ионы цинка способны и без ферментов в ограниченной степени катализировать образование комплементарных матрице олигонуклеотидов из активированных субстратов.

ДНК-полимеразы безразличны к последовательности нуклеотидов матрицы; задача этих ферментов — снять точную копию с матрицы, с какой — неважно.

Точность синтеза ДНК и коррекция

Нормальное размножение клеток требует высокой точности копирования ДНК-матрицы. Генетический материал живых организмов имеет огромные размеры. Даже у бактерий ДНК-полимераза должна практически безошибочно скопировать молекулу ДНК длиной около $3 \cdot 10^6$ п. н. Оказывается, у всех организмов точность работы репликативной машины (включающей не только ДНК-полимеразы, но и другие белки; см. ниже) как раз такова, чтобы обеспечить безошибочное воспроизведение всего генома или допустить лишь малое число ошибок. Так, у бактерий ошибки синтеза ДНК происходят не чаще чем один раз на много миллионов нуклеотидов. Молекулярные взаимодействия, на которых основаны ферментативные реакции, в частности синтез ДНК, не могут быть абсолютно надежными, кроме того, точность процесса связана с его скоростью. Для того чтобы обеспечить высокую точность наряду с высокой скоростью репликации, природе пришлось прибегнуть к специальным механизмам, один из которых — *механизм коррекции*.

ДНК-полимеразы проверяют комплементарность каждого нуклеотида матрице дважды: один раз перед включением его в состав растущей цепи и второй раз перед тем, как включить следующий нуклеотид. Очередная фосфодиэфирная связь образуется лишь в том случае, если последний (3'-концевой) нуклеотид затравки комплементарен матрице. Если же на предыдущей стадии полимеризации произошла ошибка (например, из-за того, что нуклеотид в момент полимеризации находился в необычной таутомерной форме), то репликация останавливается до тех пор, пока неправильный нуклеотид не будет удален. Некоторые ДНК-полимеразы обладают не только полимеризующей, но и 3'-экзонуклеазной активностью, которая отщепляет не спаренный с матрицей нуклеотид затравки, после чего полимеризация восстанавливается. Этот механизм, коррекция, заметно увеличивает точность работы ДНК-полимераз. Мутации, нарушающие 3'-экзонуклеазную активность ДНК-полимеразы, существенно повышают частоту возникновения прочих мутаций. Напротив, мутации, приводящие к усилению экзонуклеазной активности относительно полимеризующей, снижают темп мутирования генетического материала.

Разные ДНК-полимеразы одного организма и ДНК-полимеразы различных организмов имеют разное строение. Иногда один поли-

пептид обладает и полимеразной, и 3'-экзонуклеазной активностями, в других случаях за эти активности ответственны разные субъединицы мультисубъединичного фермента. У некоторых ДНК-полимераз корректирующая экзонуклеазная активность не обнаружена. Не исключено, что за коррекцию в этих случаях ответствен отдельный белок.

ДНК-полимеразы бактерий

Репликация лучше всего изучена у *E. coli*. У этого организма есть три ДНК-полимеразы (I, II и III). Все три фермента и их аналоги из других бактерий обладают корректирующей экзонуклеазной активностью.

ДНК-полимераза I состоит из одного полипептида длиной 911 аминокислотных остатков (а. а.) ($M_r=102\,000$ D). Этот фермент отличается от прочих ДНК-полимераз *E. coli* наличием 5'-экзонуклеазной активности. Фактически ДНК-полимераза I — это два фермента на одной полипептидной цепи: ограниченный протеолиз расщепляет эту ДНК-полимеразу на большой и малый фрагменты с разными активностями. Большой субфрагмент ДНК-полимеразы I (называемый также ДНК-полимеразой Кленова или фрагментом Кленова) обладает полимеризующей и 3'-экзонуклеазной (корректирующей) активностями. Малый субфрагмент несет 5'-экзонуклеазную активность. 5'-экзонуклеаза ДНК-полимеразы I действует на 5'-конец полинуклеотидной цепи только в составе дуплекса и отщепляет от него как моно-, так и олигонуклеотиды. Направление действия 5'-экзонуклеазы совпадает с направлением полимеризации новой цепи ДНК, т. е. в ходе полимеризации экзонуклеаза «расчищает дорогу» для полимеразы (рис. 29). Подобные свойства ДНК-полимеразы I соответствуют ее функциям в клетке: эта полимераз удаляет различного рода дефекты из ДНК в ходе репарации и слу-

жит вспомогательной полимеразой при репликации ДНК, удаляя РНК-затравки (см. следующий раздел). В клетке *E. coli* имеется несколько сотен молекул ДНК-полимеразы I.

ДНК-полимераза II — полипептид с M_r около 120 000, обладающий полимеразной и 3'-экзонуклеазной активностями. Его содержание в бактериальной клетке в несколько раз ниже, чем ДНК-

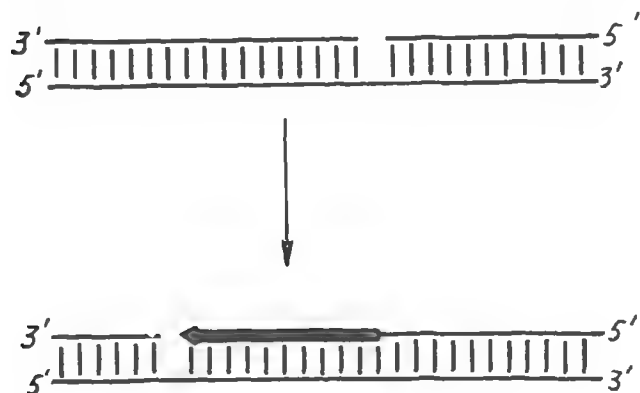


Рис. 29. Реакция переноса одноцепочечного разрыва («ник-трансляция»), осуществляемая совместным действием 5'-экзонуклеазной и полимеризующей активностей ДНК-полимеразы I *E. coli*.

полимеразы I. Эта полимеразa лучше всего работает на двуцепочечной ДНК с одноцепочечными брешами в несколько десятков нуклеотидов длиной. На этом основании можно предположить, что ДНК-полимеразa II участвует в репарации ДНК.

Главную роль в репликации ДНК у *E. coli* играет большой мультисубъединичный фермент — ДНК-полимеразa III. В клетке всего несколько таких мультимеров, приблизительно столько же, сколько репликативных вилок (см. раздел 5 этой главы). ДНК-полимеразa III

Таблица 2. Компоненты холофермента ДНК-полимеразы III *E. coli*

Субъединица	Ген	Масса, кД	Агрегаты субъединиц*
α	<i>dna E</i> (<i>pol C</i>)	140	} } } кор-фермент
ϵ	<i>dna Q</i>	25	
θ	?	10	
τ	?	83	} } ДНК-полимеразa III*
γ	<i>dna Z</i>	52	
δ	<i>dna X</i>	32	} } кополимеразa III*
β	<i>dna N</i>	37	

* Различным неполным агрегатам субъединиц ранее приписывали функцию ДНК-полимеразы III. Обнаружение этих неполных агрегатов есть следствие нестабильности холофермента в ходе выделения из клетки.

состоит из 7 типов субъединиц (табл. 2) и имеет суммарную молекулярную массу около 500 кД. Полимеразной активностью *in vitro* обладают различные неполные агрегаты субъединиц (табл. 2), но репликацию проводит полная форма фермента — холофермент, содержащий все субъединицы. Полимеразная активность присуща α -субъединице, 3'-экзонуклеазная — ϵ -субъединице, β -субъединица имеет АТРазную активность, роль которой еще будет обсуждена. Отличительная черта холофермента ДНК-полимеразы III — исключительно высокая процессивность синтеза ДНК (количество нуклеотидов, присоединяемых ферментом к растущей цепи между двумя последовательными актами диссоциации с матрицы). В табл. 3 сравнивается процессивность разных ДНК-полимераз. Видно, что по крайней мере одна из ролей многочисленных субъединиц ДНК-полимеразы III состоит в увеличении процессивности.

Для образования комплекса холофермента с матрицей-затравкой необходим АТР. Сначала АТР образует комплекс с холоферментом (с β -субъединицей), затем активированный холофермент связывается с матрицей-затравкой, что сопровождается гидролизом АТР

Таблица 3. Процессивность ДНК-полимераз

ДНК-полимераза	Матрица	Процессивность
ДНК-полимераза I	Олиго d (T)-поли d (A)	11—13
ДНК-полимераза I	Поли d (AT)	170—200
ДНК-полимераза III, кор	Однонитчатая ДНК с затравкой	10—15
ДНК-полимераза III'	То же	40
ДНК-полимераза III *	»	190
ДНК-полимераза III, холофермент	»	>5000

до ADP и фосфата. Образованный таким образом комплекс характеризуется практически неограниченной процессивностью синтеза. Видимо АТР обеспечивает необратимость присоединения к матрице (до конца копирования). Для элонгации (удлинения затравки) тоже необходим АТР, но лишь в качестве аллостерического эффектора (на этой стадии его можно заменить негидролизуемым аналогом), позволяющего ДНК-полимеразе «чувствовать» состояние энергетического баланса клетки и проводить репликацию лишь при условии достаточного энергообеспечения. При оптимальных условиях скорость синтеза ДНК холоферментом ДНК-полимеразы III *in vitro* составляет около 1000 нуклеотидов в секунду, что соответствует скорости репликации *in vivo*.

ДНК-полимеразы эукариот

В клетках эукариотических организмов обнаружены четыре ДНК-полимеразы: α , β , γ и δ . ДНК-полимераза α считается основным ферментом ядерной репликации. Содержание этого фермента заметно возрастает во время S-фазы клеточного цикла, когда происходит активный синтез ДНК. Только эта ДНК-полимераза подавляется афидиколином — ингибитором синтеза ДНК эукариот. Фермент состоит из нескольких субъединиц разного размера. Например, у дрозофилы молекулярные массы субъединиц составляют 148, 58, 46 и 42 кД. Полимеразная активность присуща самой большой из субъединиц. Молекулярная масса нативной эукариотической ДНК-полимеразы α составляет около 500 кД. Так же как в случае ДНК-полимеразы III *E. coli*, эффективность и высокая процессивность работы полимеразы α зависят до дополнительных субъединиц, которые сами по себе полимеризующей активностью не обладают. Одна из субъединиц ДНК-полимеразы α оказалась ДНК-праймазой — ферментом, необходимым для инициации новых цепей ДНК (см. ниже); ассоциация с праймазой не характерна для ДНК-полимераз бактерий.

ДНК-полимераза β состоит из одного полипептида размером около 40 кД. Действие этого фермента, видимо, не процессивно. Наибольшая активность наблюдается на двуцепочечной ДНК с короткими брешами. Можно думать, что основная роль этой полимеразы состоит в репарации ДНК.

ДНК-полимераза γ , по-видимому, состоит из нескольких идентичных субъединиц с молекулярной массой около 50 кД. Это митохондриальный фермент, который осуществляет синтез ДНК митохондрий. Сходный фермент обнаружен в хлоропластах растений.

ДНК-полимераза δ имеет молекулярную массу около 150 кД и состоит из двух неидентичных субъединиц. Этот фермент обладает сопряженной 3'-экзонуклеазной активностью. Его роль в синтезе ДНК не ясна.

У низших эукариот дрожжей обнаружены две ДНК-полимеразы: I и II. По своей роли в репликации ДНК они аналогичны ДНК-полимеразам III и I *E. coli* соответственно (см. ниже). У полимеразы II дрожжей есть ассоциированная 3'-экзонуклеазная активность.

2. ОСНОВНЫЕ ПРИНЦИПЫ РЕПЛИКАЦИИ

Инициация цепей ДНК

ДНК-полимеразы не способны инициировать новые цепи ДНК. Они могут лишь достраивать уже имеющуюся *затравку*. Это свойство, по-видимому, связано с предъявляемым к синтезу ДНК требованием высокой точности. Как уже говорилось, безошибочная работа ДНК-полимераз в существенной степени определяется их неспособностью удлинять затравку, последний нуклеотид которой не спарен с матрицей. Так как свободные нуклеотиды с матрицей не спариваются, то соединение самого первого нуклеотида цепи ДНК со следующим аналогично присоединению нуклеотида к не спаренному с матрицей концу затравки и оказывается невозможным. Как же, в таком случае, начинается синтез цепей ДНК?

Оказывается, новосинтезированные цепи ДНК всегда содержат на 5'-конце несколько рибонуклеотидов. Иными словами, синтез ДНК начинается с синтеза РНК. РНК-затравку для синтеза ДНК образует специальный фермент, называемый ДНК-праймазой (от англ. праймер — затравка). Праймаза может быть отдельным ферментом, как у бактерий, или входить в качестве субъединицы в ДНК-полимеразу (как у ДНК-полимеразы α животных). В любом случае праймаза — это фермент, отличный от РНК-полимераз, синтезирующих разнообразные клеточные РНК и тоже способных инициировать синтез новых полинуклеотидных цепей (см. гл. VII). Почему же в таком случае для инициации цепей ДНК используются рибонуклеотиды? Возможное объяснение состоит в том, что в ходе эволюции праймазы произошли из РНК-полимераз. Но есть и другое, функциональное, объяснение. Поскольку требования инициа-

ции и безошибочности синтеза противоречат друг другу, затравку, синтезированную иницирующим (а значит, неточным) ферментом, необходимо как-то отличить от остальной новосинтезированной цепи. Можно думать, что именно поэтому она и состоит из рибонуклеотидов. После того как цепь ДНК начала синтезироваться, РНК-затравки удаляются и образовавшиеся бреши застраиваются ДНК-полимеразой, т. е. с высокой точностью.

Расплетание двойной спирали ДНК в ходе репликации

Нативные ДНК двуспиральны; следовательно, перед репликацией цепи родительской молекулы, матричные цепи ДНК, должны быть разделены. Эту реакцию осуществляют два типа белков: *хеликазы* и

SSB-белки (от англ. single strand binding — белки, связывающиеся с однонитевой ДНК). Хеликазами называют ДНК-зависимые АТРа-зы, использующие энергию гидролиза АТФ для расплетания двойной спирали (helix) ДНК. Считается, что хеликаза, движимая гидролизом АТФ, однонаправленно «едет» по одной из цепей ДНК, расплетая перед собой двойную спираль. Есть хеликазы, которые едут от 5'-конца к 3'-концу цепи ДНК, и есть другие, перемещающиеся в обратном направлении. В результате работы хеликаз возникает «вилка» из двуспирального участка ДНК и двух одноцепочечных ветвей. Ренатурации одноцепочечных участков ДНК препятствует их связывание SSB-белком, имеющим избирательное сродство к однонитевой ДНК (рис. 30). SSB-белки и хеликазы обнаружены у многих про- и эукариотических организмов. Роль SSB-белка в репликации, по-видимому, состоит в том, чтобы расправить ДНК, вытянуть ее и удалить возможные элементы вторичной структуры, которые могли бы образоваться в самокомплементарных участках ДНК. Связывание одноцепочечной ДНК с SSB-белком стимулирует ДНК-

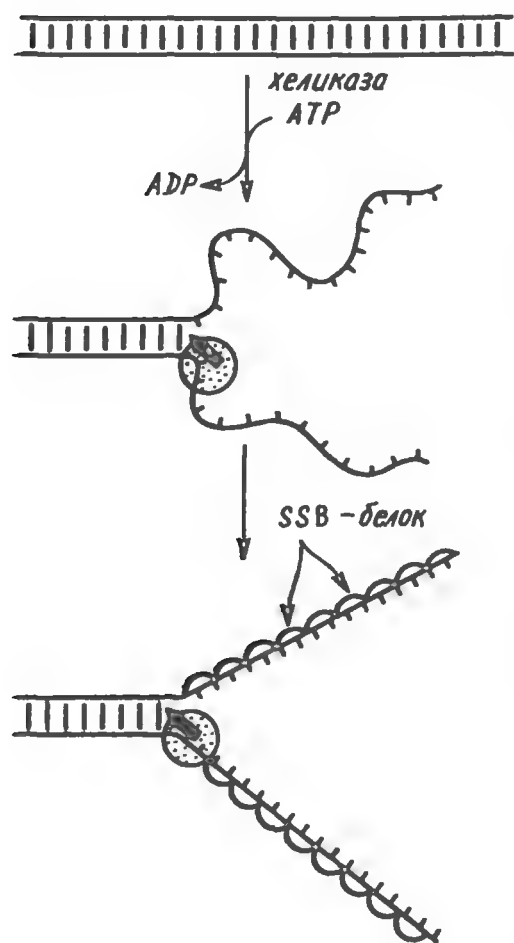


Рис. 30. Расплетание двойной спирали ДНК хеликазой и SSB-белком

(Существуют хеликазы, перемещающиеся по цепи ДНК в направлении 5' → 3', и хеликазы, перемещающиеся в обратном направлении. SSB-белок связывает образующуюся в результате действия хеликазы одноцепочечную ДНК и препятствует ее ренатурации.)

полимеразу и повышает точность ее работы. Этот эффект вызывается не только разрушением вторичной структуры одноцепочечной ДНК, но и непосредственным взаимодействием ДНК-полимеразы с SSB-белком, поскольку обычно полимеразу стимулирует лишь «свой» SSB-белок, но не аналогичный белок из другого источника. SSB-белок *E. coli* — тетрамер, состоящий из идентичных субъединиц размером 19 кД. SSB-белок связывается с ДНК кооперативно, т. е. за счет белок-белковых взаимодействий тетрамеры покрывают ДНК вплотную друг к другу.

Прерывистый синтез ДНК

Так как цепи ДНК в дуплексе антипараллельны, то очевидно, что направление расплетания двойной спирали при репликации совпадает с направлением синтеза ДНК лишь для одной матричной цепи, но противоположно направлению синтеза ДНК на комплементарной матрице (рис. 31). Это значит, что лишь на одной из матричных цепей синтез ДНК может происходить непрерывно. Как синтезируется ДНК на второй матрице? Показано, что ДНК синтезируется сравнительно короткими фрагментами, называемыми *фрагментами*

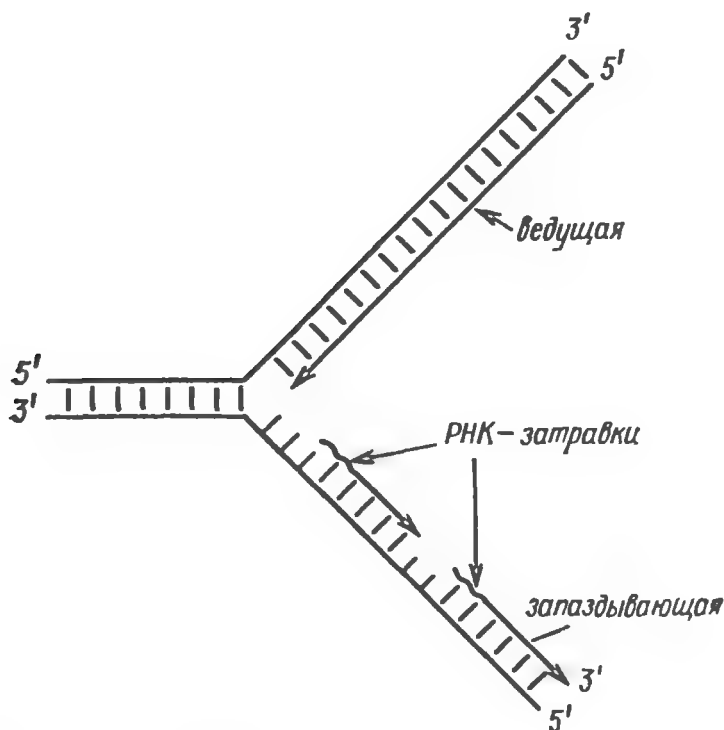


Рис. 31. Репликативная вилка

Направление синтеза ДНК совпадает с направлением расплетания исходной двойной спирали лишь для одной из новосинтезированных цепей (ведущая цепь). Вторая цепь синтезируется прерывисто, сравнительно короткими фрагментами. Инициация каждого такого фрагмента Оказаки происходит лишь после того, как образуются достаточно протяженный одионитовой участок матрицы

Оказаки. Таким образом, синтез ДНК на двух матричных цепях исходной молекулы заметно различается. Новосинтезированная цепь, которая синтезируется непрерывно, называется *ведущей* (англ. *leading*), другая цепь называется *запаздывающей* (англ. *lagging*). Каждый фрагмент Оказаки имеет на 5'-конце несколько рибонуклеотидов — результат действия праймазы. Характерный размер фрагментов Оказаки различается для бактерий и эукариот: у бактерий они имеют длину около 1000 нуклеотидов, у эукариот они короче, порядка 100 нуклеотидов. Через некоторое время после синтеза РНК-затравки удаляются, бреши застраиваются ДНК-полимеразой, а фрагменты сшиваются в одну ковалентно-непрерывную цепь ДНК предназначенным специально для этого ферментом, ДНК-лигазой.

ДНК-лигазы обнаружены у самых разных организмов. Они нуждаются в высокоэнергетических кофакторах: например, лигаза *E. coli* нуждается в NAD^+ , а широко используемая для генно-инженерных методик лигаза фага Т4 — в АТФ. Прежде чем образовать фосфодиэфирную связь, соединяющую два фрагмента ДНК, лигаза активирует 5'-конец одного из фрагментов с помощью высокоэнергетического кофактора (рис. 32).

3. РЕПЛИКАЦИЯ ДНК *E. coli*

Детальную картину репликации ДНК лучше всего рассмотреть на примере *E. coli*. Описанные в настоящем разделе механизмы репликации справедливы по крайней мере для других бактерий.

Для выяснения механизма репликации бактериальной хромосомы незаменимую роль сыграл анализ разнообразных мутантов, нарушающих репликацию ДНК. Синтез ДНК — функция жизненно важная, и мутации, инактивирующие ферменты синтеза ДНК, летальны. Поэтому, как и в других подобных случаях, были использованы условно летальные мутации, в частности температурочувствительные (*ts*).

В табл. 4 приведены данные о белках, участвующих в репликации ДНК *E. coli*, и о кодирующих их генах. Кроме белков, действие которых уже обсуждено, есть ряд белков, о которых речь еще не шла.

Основной хеликазой при репликации ДНК *E. coli* служит комплекс из шести молекул белка-продукта гена *dnaC* и шести молекул белка-продукта гена *dnaB* (собственно хеликазой является DnaB-белок). С этой хеликазой связана праймаза, кодируемая у *E. coli* геном *dnaG*. Такой комплекс из хеликазы и праймазы называют *праймосомой*. Гидролизую АТФ, праймосома едет по матрице для синтеза запаздывающей цепи ДНК в направлении от 5'- к 3'-концу и расплетает перед собой двойную спираль, а за собой оставляет РНК-затравки для синтеза фрагментов Оказаки. Репликацию ДНК *E. coli* можно воспроизвести *in vitro* с праймосомой, состоящей только из белков DnaB, DnaC и DnaG, но в клетке праймосома может

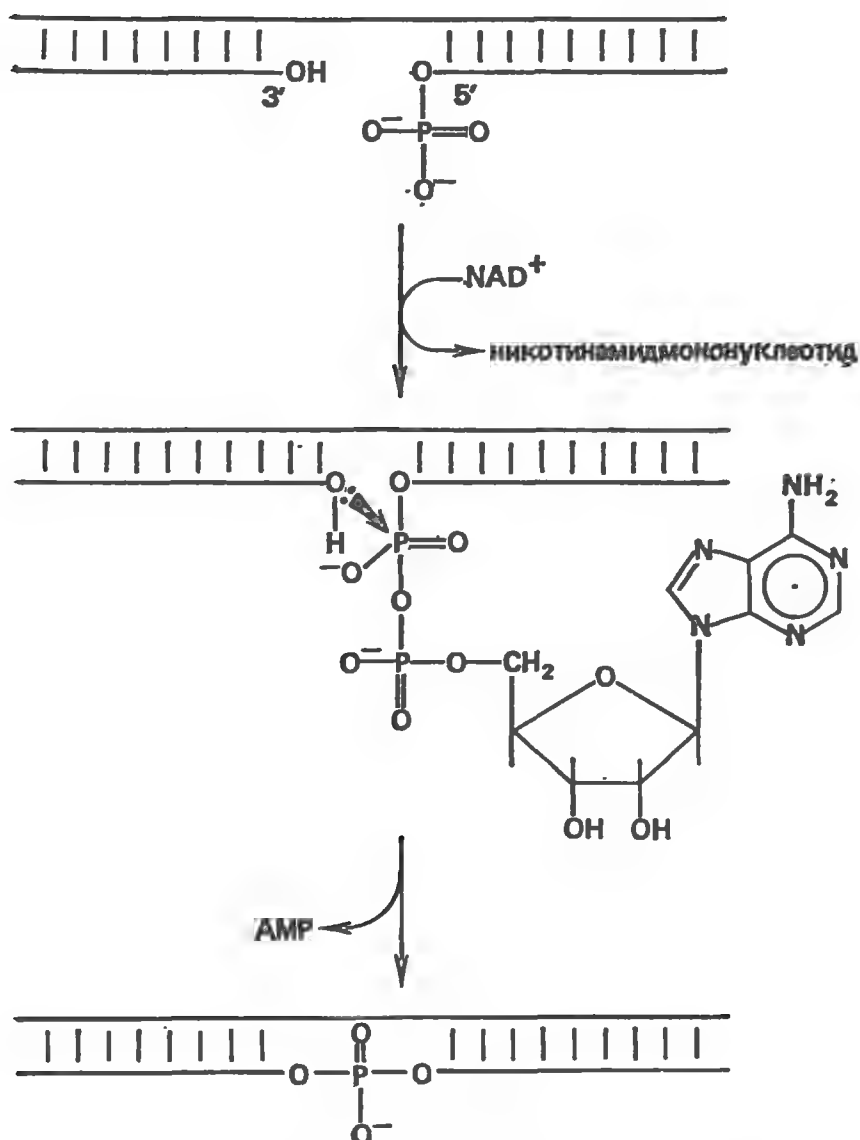


Рис. 32. Механизм действия ДНК-лигазы *E. coli*

(Лигаза аденилирует 5'-конец одноцепочечного разрыва ДНК и тем самым активирует его. В результате нуклеофильной атаки 3'-концевой ОН-группы на ближайший фосфат образуется фосфодиэфирная связь и освобождается АМР.)

быть устроена сложнее; по крайней мере в некоторых модельных системах в ее состав входят еще четыре белка: π , π' , π'' и i . Функции этих белков неясны, лишь для белка π' известно, что он является ДНК-зависимой АТРазой и может продвигать праймосому в направлении 5' → 3' по однонитевой ДНК, покрытой SSB-белком. Состав входящих в праймосому белков, видимо, определяется механизмом сборки праймосомы в ходе инициации раунда репликации (см. раздел 5 этой главы).

Вспомогательной хеликазой при синтезе ДНК *E. coli* служит белок — продукт гена *rep*. Рер-хеликаза едет по матрице для синтеза ведущей цепи в направлении от 3'-к 5'-концу. Таким образом, хеликазы DnaB и Rep едут по разным цепям ДНК, но в силу ан-

Таблица 4. Белки репликации *E. coli* и их гены

Ген	Белок	Агрегат	Функция
<i>dnaE</i> <i>dnaN</i> <i>dnaZ</i> <i>dnaX</i> <i>dnaQ</i> ? ? <i>dnaB</i> <i>dnaC</i> <i>dnaG</i> <i>dnaT</i> ? ? ?	α-Субъединица β-Субъединица γ-Субъединица δ-Субъединица ε-Субъединица θ-Субъединица τ-Субъединица Хеликаза DnaC Праймаза i п п' п''	ДНК-полимера- за III, холофер- мент Белки, участвую- щие в сборке и функ- ционировании прай- мосомы	Элонгация цепей ДНК, основной фермент репли- кации Расплетание двойной спирали и синтез РНК- затравок на запаздываю- щей цепи
<i>polA</i>	ДНК-полимераза I		Удаление РНК-затра- вок, застройка брешей
<i>lig</i>	ДНК-лигаза		Сшивание фрагментов
<i>ssb</i>	SSB		Оказаки Связывание одоните- вой ДНК
<i>hupA</i> <i>hupB</i> <i>rep</i>	} HU Rep		Гистоноподобный белок
<i>gyrA</i> <i>gyrB</i>	А-Субъединица } В-Субъединица }	ДНК-гираза	Расплетание двойной спирали
<i>topA</i>	ω, топоизомераза I		Снятие топологических затруднений при репли- кации кольцевых ДНК, инициация раунда репли- кации
<i>rnh</i>	РНКаза H		Обеспечение избира- тельности инициации реп- ликации
<i>dnaA</i>	DnaA		Участие в инициации репликации и обеспечение ее избирательности Инициация раунда реп- ликации

типараллельности цепей в двойной спирали в одну сторону, в сто-
рону раскрытия репликативной вилки.

Теперь можно нарисовать репликативную вилку со всеми дей-
ствующими там белками (рис. 33). Дуплекс родительской молекулы
расплетают две хеликазы — Rep и DnaB — в составе праймосомы.
Образующиеся одноцепочечные участки кооперативно покрывает
SSB-белок. Холофермент ДНК-полимеразы III едет по одной из
матричных цепей в направлении раскрытия вилки и синтезирует
ведущую цепь ДНК. По другой матричной цепи в том же направле-
нии едет праймосома. Время от времени входящая в состав праймо-

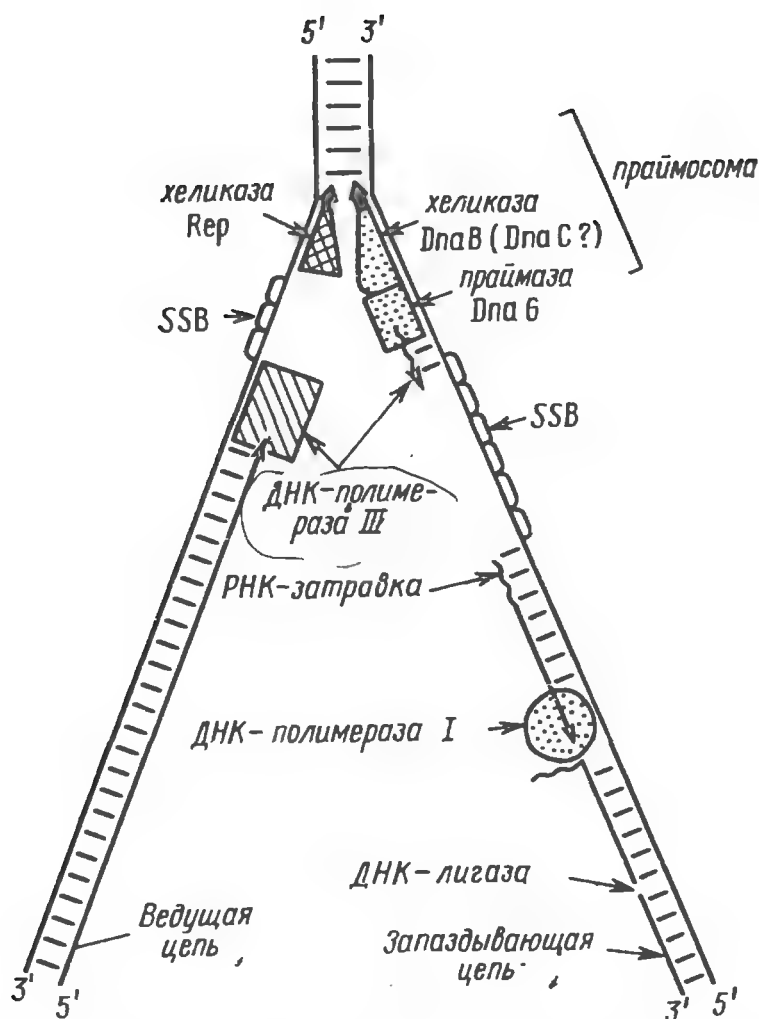


Рис. 33. Схема синтеза ДНК в репликативной вилке

сомы праймаза, белок DnaG, синтезирует РНК-затравки для запаздывающей цепи. Вторая молекула холофермента ДНК-полимеразы III удлиняет эти затравки до тех пор, пока не упрется в предыдущую затравку, т. е. синтезирует фрагменты Оказаки. Затем действует ДНК-полимераза I, которая продолжает удлинять фрагменты Оказаки, одновременно гидролизую РНК-затравку предыдущего фрагмента, используя свою 5'-экзонуклеазную активность. После действия ДНК-полимеразы I между двумя соседними фрагментами остается только одноцепочечный разрыв, который зашивает ДНК-лигаза. Таким образом, в репликативной вилке одновременно работают около 20 разных полипептидов, осуществляя сложный, высокоупорядоченный и энергоемкий процесс. Не говоря уже о том, что каждый нуклеотид переходит в ДНК из богатого энергией предшественника, множество молекул АТФ тратится на действие хеликаз, на синтез РНК-затравок, которые затем удаляются, на активацию ДНК-полимеразы III при переходе на каждый новый фрагмент Оказаки запаздывающей цепи и на работу топоизомераз по раскручиванию взаимозакрученных цепей ДНК (см. ниже). Такая цена высокой точности и скорости репликации.

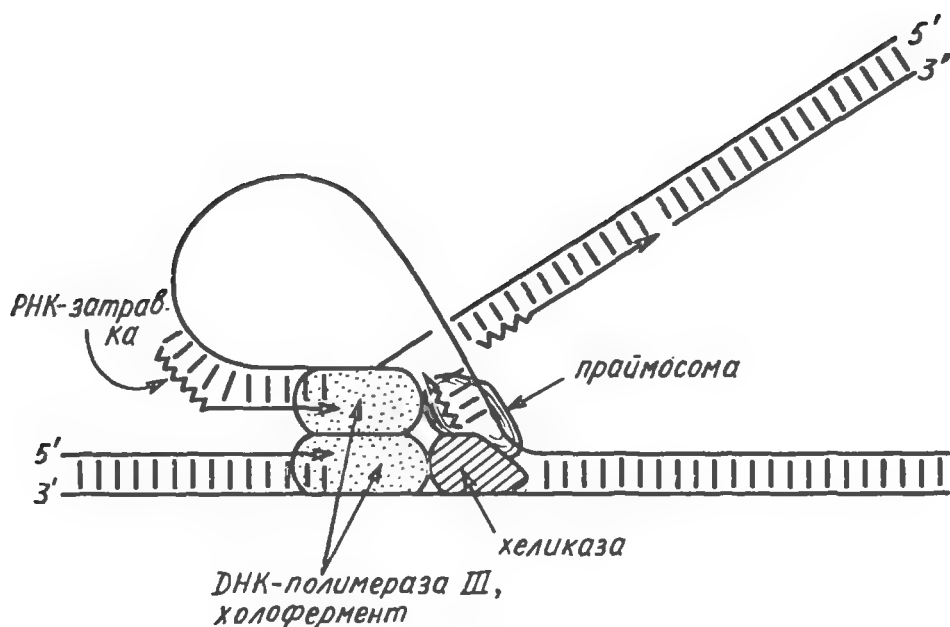


Рис. 34. Схема возможной организации репликативной вилки

Две молекулы ДНК-полимеразы одновременно и синхронно синтезируют и ведущую, и запаздывающую цепи ДНК. Когда полимераза, синтезирующая запаздывающую цепь, доходит до начала предыдущего фрагмента Оказаки, она «перескакивает» на 3'-конец затравки, которую синтезирует в этот момент праймосома; по всей вероятности это событие требует гидролиза АТФ. Огромный белковый комплекс, осуществляющий одновременный синтез ДНК на обеих матричных цепях репликативной вилки, предложено называть «реплисомой», ориентировочная масса комплекса 1400.)

Холофермент ДНК-полимеразы III может димеризоваться. На этом основании выдвинуто предположение, что две молекулы холофермента, праймосома, и, возможно, дополнительная хеликаза образуют в клетке единый комплекс — *реписому*, которая движется по ДНК, одновременно синтезируя обе новые цепи (рис. 34). Не исключено, что в состав гипотетической реписомы входят и некоторые ферменты биосинтеза нуклеотидов — предшественников ДНК. Это было бы разумно, учитывая, что дезоксирибонуклеотиды нужны в клетке только для синтеза ДНК.

Нарисованную картину бактериальной репликации подтверждает фенотип мутантов с нарушением синтеза ДНК. *ts*-Мутации в генах *dnaE*, *dnaB*, *dnaG*, *dnaZ*, *dnaX* и *dnaN* сразу останавливают репликацию при переносе клеток на непермиссивную температуру. Мутации в генах *polA* и *lig* приводят к тому, что в клетке накапливается большое количество несшитых фрагментов Оказаки. Мутации в генах *dnaA*, *topA* и *rnh* влияют на инициацию цикла репликации, которая будет рассмотрена ниже.

Репликацию ДНК *E. coli* удалось воссоздать *in vitro* в системе из очищенных белков. Для реакции необходимы все вышеперечисленные белки, а для оптимального синтеза — белок HU — гистоноподобный белок *E. coli*. Нужны также белки, участвующие в инициации репликации, и топоизомеразы.

4. ТОПОЛОГИЧЕСКИЕ ПРОБЛЕМЫ РЕПЛИКАЦИИ

До сих пор никак не учитывался тот факт, что комплементарные цепи ДНК закручены друг вокруг друга в спираль. Между тем это существенно. Большинство молекул ДНК бактерий и некоторые ДНК эукариот являются кольцевыми. Из-за спиральной закрученности цепи этих молекул оказываются *зацепленными* — их невозможно разделить, не порвав хотя бы одну из них. Даже если бы цепи не были зацепленными (т. е. ДНК не была бы кольцевой), при скорости движения репликативной вилки 1000 н. п. в секунду вся непрореплицированная часть ДНК должна вращаться со ско-

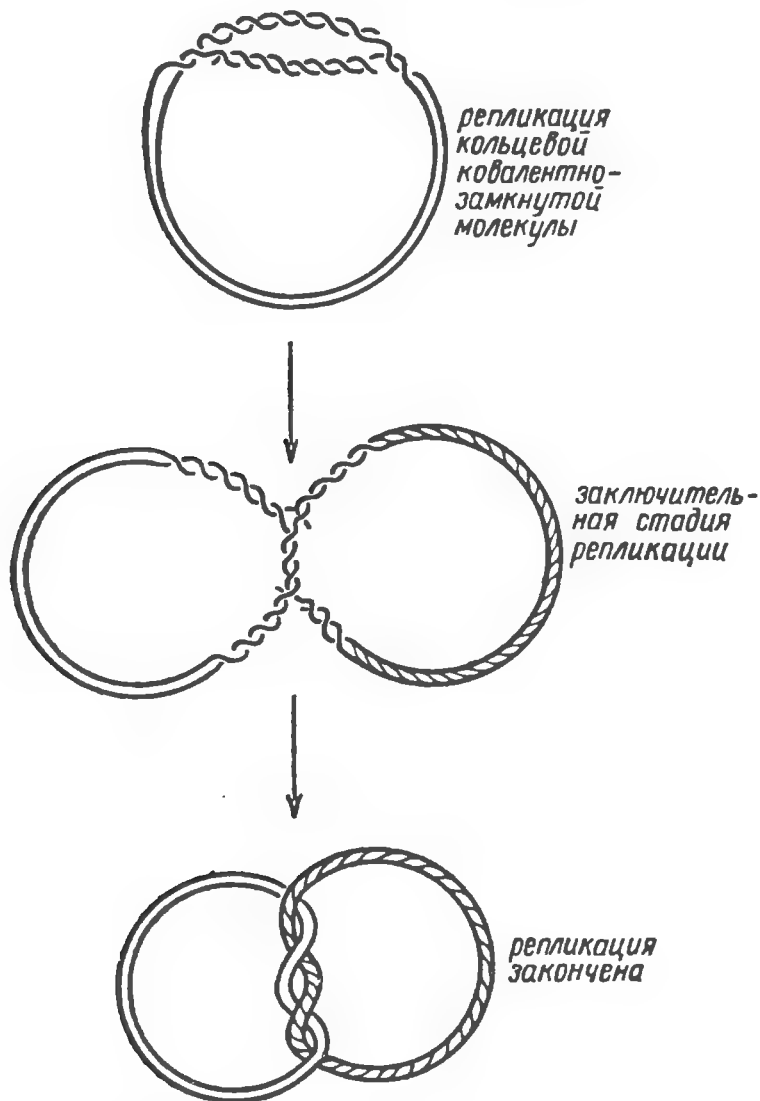


Рис. 35. Образование катенанов при репликации кольцевых ДНК

[Взаимозакрученность двух матричных цепей в кольцевых молекулах ДНК часто приводит к образованию в результате репликации зацепленных друг за друга дочерних молекул (катенанов). Расцепить катенаны способна топоизомераза второго типа (см. гл. 1)]

ростью 6000 оборотов в минуту! Все эти проблемы разрешаются присутствием в клетке *топоизомераз* (см. гл. I). Топоизомеразы на время вносят в зацепленные друг за друга цепи кольцевых молекул разрывы, которые необходимы для их разделения, т. е. выступают в роли «шарниров», позволяющих цепям ДНК раскрутиться. Способность бактериальной топоизомеразы II — ДНК-гиразы — отрицательно сверхспирализовать ДНК в АТР-зависимой реакции не только снимает вопрос о вращении всей непрореплицировавшейся части ДНК, но и облегчает действие хеликаз, поскольку отрицательная сверхспирализация, которую гираза создает перед вилкой, способствует расплетанию ДНК.

На заключительной стадии репликации кольцевых молекул часто остается одно или несколько зацеплений цепей исходной молекулы друг за друга. Это приводит к тому, что двуцепочечные кольца дочерних молекул также оказываются зацепленными, образуют *катенан* (рис. 35). ДНК-гираза может расцепить зацепленные кольца, используя свою способность вносить временный двуцепочечный разрыв. Такая активность гиразы действительно существенна для репликации ДНК, поскольку в мутантах по гиразе на непермиссивной температуре наблюдается нерасхождение дочерних молекул кольцевых ДНК после репликации. Важно отметить, что топоизомеразы необходимы для завершения репликации не только кольцевых молекул, но и очень длинных линейных эукариотических хромосом: две очень длинные дочерние молекулы не могут разойтись достаточно быстро, поскольку после репликации оказываются запутанными подобно катенанам, образующимся на заключительной стадии репликации кольцевых ДНК. Действительно, мутанты эукариот (дрожжей) с нарушенной топоизомеразой II дефектны по расхождению дочерних хромосом в митозе.

По крайней мере у бактерий топоизомеразы играют важную роль не только в ходе репликации и на завершающих стадиях этого процесса, но и при инициации раунда репликации.

5. ИНИЦИАЦИЯ РАУНДА РЕПЛИКАЦИИ

Начавшийся процесс репликации хромосомы бактерии продолжается до тех пор, пока не удвоится вся ДНК. В этом смысле бактериальная хромосома представляет собой единицу репликации — *репликон*. Другие молекулы ДНК, которые могут присутствовать в бактериальных клетках (см. гл. V), также представляют собой отдельные репликоны.

Репликация каждого бактериального репликона, в частности хромосомы *E. coli*, как правило, начинается в одной избранной области ДНК, называемой *ориджином* репликации (от англ. origin — начало, обозначается *ori*). Ориджин репликации каждого репликона имеет вполне определенную последовательность ДНК. В результате инициации раунда репликации на ориджине образуются одна или

две репликативные вилки. Если возникает одна репликативная вилка, как, например, у плазмиды ColE1, то, объехав всю кольцевую молекулу, она заканчивает синтез на ориджине. Такая репликация называется *однонаправленной*. На многих ориджинах, в частности на хромосомном ориджине *E. coli*, инициируется *двунаправленная* репликация, т. е. возникают две репликативные вилки, перемещающиеся в противоположных направлениях до тех пор, пока не встретятся. Реплицирующиеся кольцевые репликоны имеют характерную структуру, напоминающую греческую букву Θ и называемую *тэта-структурой* (см. рис. 10, верх).

Последовательность ориджина способствует необходимому для начала синтеза ДНК расплетанию двойной спирали ДНК и служит участком сборки, «посадки» на ДНК активного комплекса белков, осуществляющих репликацию. Чем же ориджины репликации отличаются от прочих последовательностей ДНК, что определяет их специфичность? Для разных репликонов ответ может быть различным, однако часто оказывается, что специфичность ориджина определяется специальным белком, участвующим в инициации синтеза ДНК и способным избирательно связываться с последовательностью нуклеотидов данного ориджина. Наличие на одном репликоне ориджина и гена, кодирующего специфичный к нему белок-инициатор, обеспечивает самоподдержание этого репликона в клетке.

Инициация репликации хромосомы E. coli

Белок-инициатор, специфически узнающий хромосомный ориджин *E. coli oriC*, кодируется геном *dnaA*. В последовательности *oriC* присутствует несколько участков связывания DnaA-белка. Молекулы белка-инициатора связываются с этими участками, если ДНК ориджина благодаря действию ДНК-гиразы сверхспирализована, и расплетают ДНК в находящейся поблизости АТ-богатой области ориджина.

Можно думать, что участие ДНК-гиразы в инициации необходимо не только для того, чтобы облегчить белку DnaA задачу расплетания ориджина, но и для проверки целостности цепей ДНК — будущих матриц для синтеза: действительно, сверхспирализовать какую-либо кольцевую ДНК (или участок ДНК, концы которого фиксированы) можно лишь в том случае, если эта ДНК не содержит разрывов.

С комплексом DnaA-белка с расплетенным *oriC* связывается DnaC-белок, после чего с образовавшимся комплексом может связаться хеликаза DnaB, которая обеспечивает дальнейшее расплетание ориджина, необходимое для сборки *реплисома* (рис. 36).

Для инициации репликации хромосомы *E. coli* необходима также транскрипция области ориджина РНК-полимеразой. Поскольку

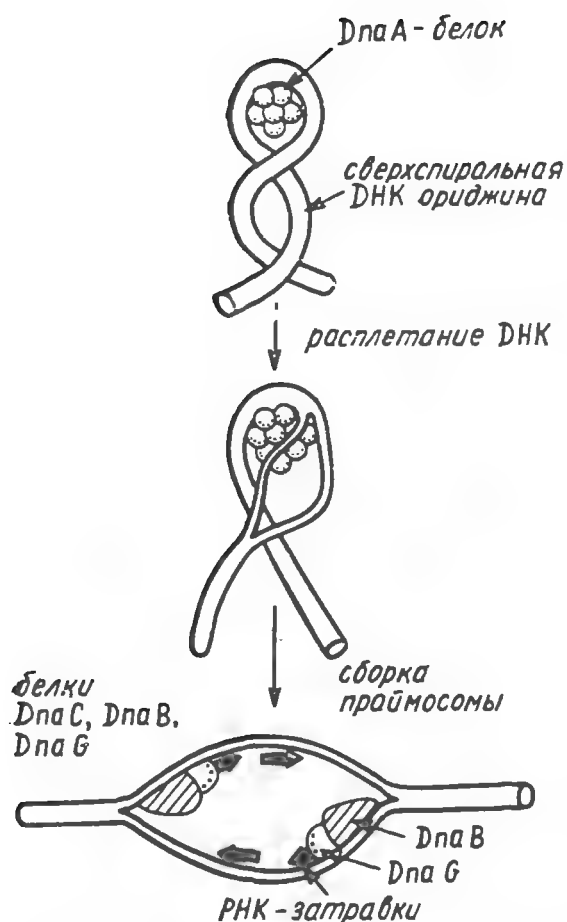


Рис. 36. Инициация репликации хромосомы *E. coli*

(Хромосомный ориджин, *oriC* специфически узнает белок-инициатор ДнаА. С комплексом ДнаА и *oriC* связывается белок ДнаС. С тройным комплексом, ДНК которого сверхспирализована, связываются ДнаВ-хеликаза и праймаза. Белок ДнаВ начинает расплетать дуплекс ДНК, а праймаза синтезирует первые затравки)

первые РНК-затравки синтезирует, по-видимому, праймаза, роль РНК-полимеразы скорее всего заключается в «транскрипционной активации» ориджина. Можно думать, что локальное расплетение ДНК, создаваемое РНК-полимеразой в ходе транскрипции (см. гл. VII), облегчает наряду с отрицательной сверхспирализацией стабильное «раскрытие» ориджина белками ДнаА, ДнаВ и ДнаС.

Два фермента обеспечивают высокую избирательность инициации синтеза ДНК, ограничивая инициацию репликации только ориджином. Это топоизомераза I и РНКазы H, избирательно гидролизующая РНК в составе гибридных дуплексов с ДНК. Действие этих ферментов направлено против гибридных ДНК—РНК-участков, которые могут случайно образоваться на ДНК при транскрипции и послужить затравками для начала синтеза ДНК. Возможная роль в этом процессе РНКазы H очевидна: она способна непосредственно гидролизовать РНК во всех таких участках. Что касается

роли топоизомеразы I, то необходимо отметить, что гибриды ДНК—РНК образуются лишь в том случае, если ДНК сверхспирализована (образование гибридного дуплекса снимает часть избыточной энергии сверхспирализации), причем сверхспирализована достаточно сильно, чтобы локальные нарушения нормальной вторичной структуры ДНК могли способствовать гибридизации с РНК. Топоизомераза I может релаксировать сверхспиральную ДНК лишь в том случае, если она сверхспирализована отрицательно и достаточно сильно, т. е. в условиях, способствующих возникновению на ДНК упомянутых локальных нарушений вторичной структуры. Таким образом, можно думать, что одна из функций этого фермента состоит в поддержании нормальной вторичной структуры ДНК, препятствующей ее гибридизации с РНК и образованию затравки. В мутантах *E. coli* по РНКазе H (ген *rnh*) или по топоизомеразе I (ген

topA) инициация репликации может происходить не только на ориджине, но и во многих других местах.

Смысл ограничения инициации репликации областью ориджина становится понятным, если вспомнить, что репликация строго регулируется (см. ниже). Мишенью различных регуляторных воздействий является ориджин, поэтому для контроля репликации необходимо подавить инициацию в прочих местах.

Другие способы инициации репликации целесообразно рассмотреть в связи с вопросом о регуляции репликации. Молекулярные основы подобной регуляции лучше всего изучены в случае некоторых бактериальных плазмид.

Регуляция репликации

При каждом клеточном делении каждая молекула ДНК должна удваиваться, т. е. на каждом ориджине должен происходить в точности один акт инициации репликации. В противном случае постепенно происходила бы утеря репликона или его бесконтрольное накопление. Более того, даже если репликон удваивается в среднем точно один раз на каждое клеточное деление, возможны существенные вариации количества копий этого репликона вокруг среднего значения в разных клетках бактериальной популяции. Такие вариации недопустимы, так как тоже в конце концов ведут к потере репликона. Таким образом, к регуляции репликации предъявляются достаточно жесткие требования: регуляторная система должна чувствовать отклонения в обе стороны от среднего числа копий данного репликона и соответствующим образом менять частоту инициации на ориджине. Очевидно, что частота инициации должна быть согласована также со скоростью роста клеток.

Некоторые механизмы регуляции репликации

К наиболее изученным репликонам принадлежат плаزمида ColE1 и родственные ей плазмиды. Ориджины плазмид этой группы были использованы при конструировании многих векторов для молекулярного клонирования. Репликация ColE1 начинается на единственном ориджине и происходит однонаправленно. Механизм инициации репликации этой плазмиды существенно отличается от механизма инициации репликации хромосомы *E. coli*. Для инициации необходимо, чтобы РНК-полимераза синтезировала на ориджине РНК-предшественник затравки, так называемую РНКII. Транскрипция РНКII начинается на расстоянии около 550 п. н. от точки начала синтеза ДНК. Пока не выясненные способности последовательности ориджина и определенные элементы вторичной структуры РНКII способствуют образованию гибридного дуплек-

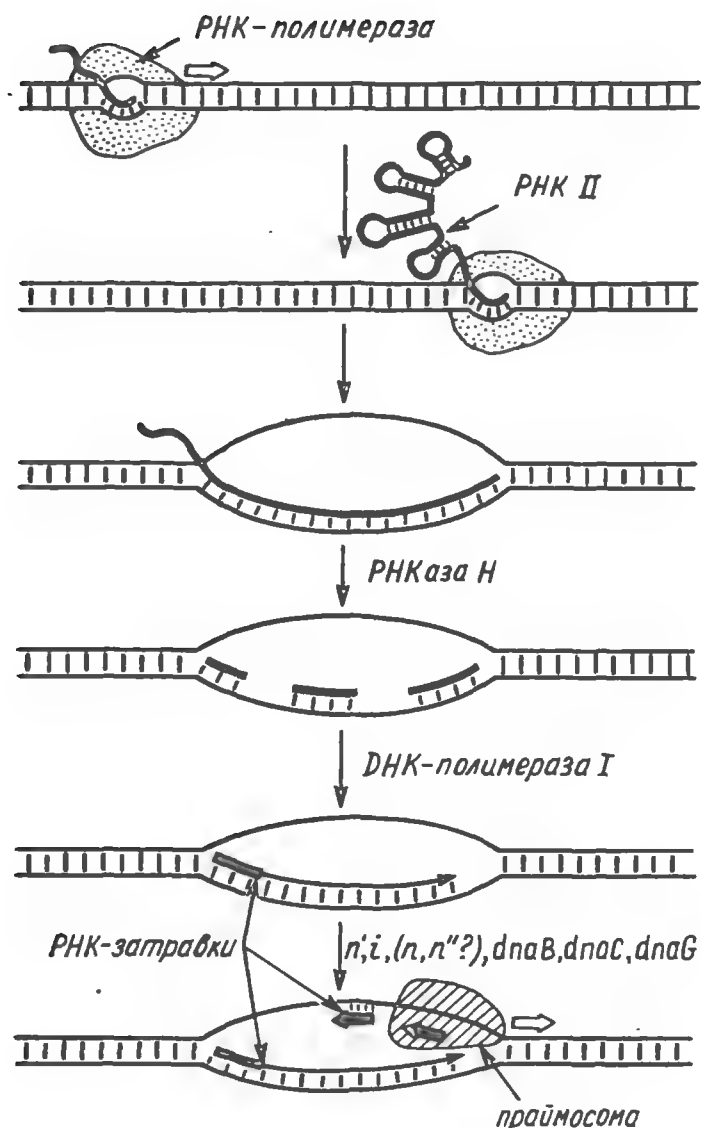


Рис. 37. Инициация репликации плазмы ColE1

са РНКII с ДНК в области ориджина (рис. 37). Для образования гибридного дуплекса необходимо, чтобы ориджин плазмиды был сверхспирализован, т. е. и в этом случае молекула ДНК готова к репликации лишь при условии ковалентной непрерывности обеих матричных цепей (в частности, репликация плазмиды не начнется, пока не закончится предыдущий раунд репликации). Образовавшийся гибрид РНК — ДНК подвергается частичному расщеплению РНКазой Н, в результате чего образуется РНК-затравка для синтеза ДНК. Для этого репликона РНКазы Н служит не только фактором специфичности, ограничивающим инициацию районом ориджина, но и непосредственно участвует в инициации. Пока не ясно, какие особенности структуры ориджина препятствуют полной деградации РНК в гибридном дуплексе и ограничивают действие РНКазы Н образованием затравки. Синтез ДНК на РНК-затравках

начинает у плазмид типа ColE1 ДНК-полимераза I; в этом состоит существенное отличие инициации этих репликонов от других. Необходимость именно в этой ДНК-полимеразе, по-видимому, связана с тем, что после действия РНКазы H не только затравка, но и другие фрагменты РНКII остаются спаренными с ДНК. Для их удаления необходима 5'-экзонуклеазная активность ДНК-полимеразы I. ДНК-полимераза I образует сравнительно короткий фрагмент ДНК длиной в несколько сотен нуклеотидов, при этом образуется так называемая D-петля, в которой одна цепь исходной плазмиды спарена с новосинтезированной цепью, а вторая вытеснена и находится в одноцепочечном состоянии (рис. 37). По всей видимости, именно на этой однонитевой ДНК собирается комплекс праймосомных белков (состоящий по крайней мере из продуктов генов *dnaB*, *dnaC*, *dnaG* и *dnaT*) и дальнейший синтез ДНК осуществляют уже ДНК-полимераза III и праймосома, т. е. возникает и продвигается в одном направлении обычная репликативная вилка (рис. 37).

Критическим событием инициации репликации плазмид этого типа является комплементарное спаривание РНКII с ДНК ориджина. Естественно, именно эта стадия инициации репликации и регулируется. Оказалось, что с области ДНК, кодирующей начало (5'-концевую часть) РНКII, считывается еще одна РНК — РНКI,

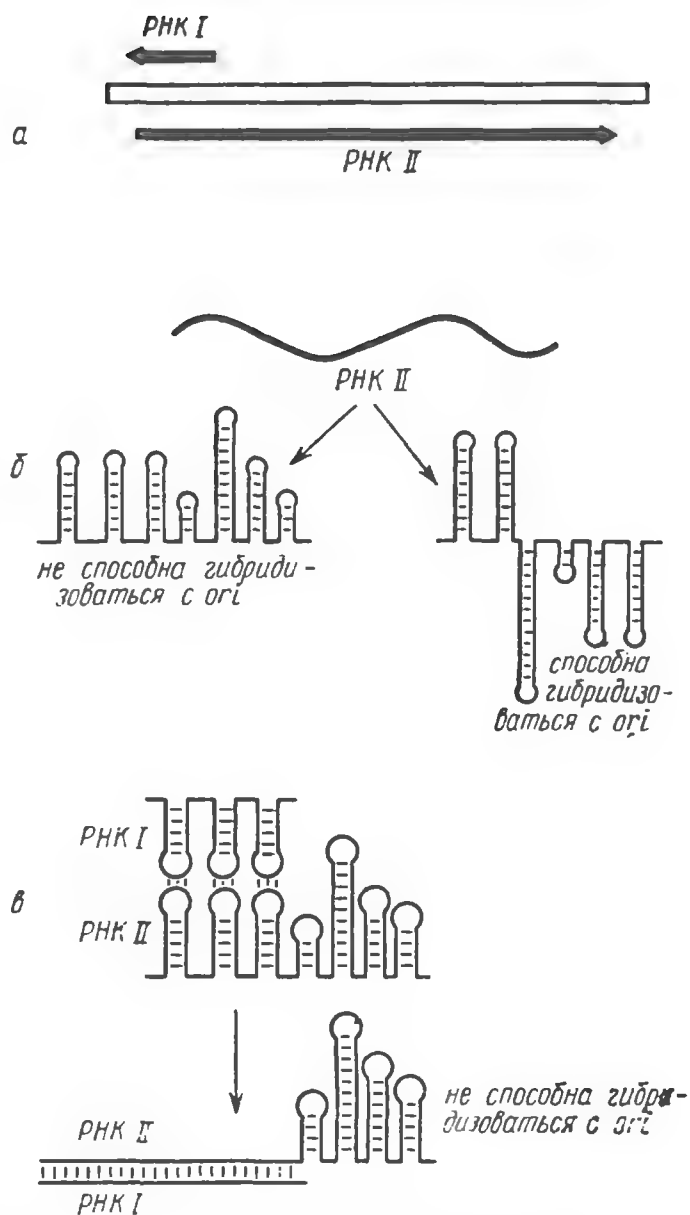


Рис. 38. Взаимодействие РНКII и РНКI:

а — РНКII и РНКI считываются с разных цепей одного участка ДНК и поэтому комплементарны. *б* — РНКII может образовывать два альтернативных варианта вторичной структуры, из которых лишь один способствует связыванию РНКII с ДНК ориджина; *в* — взаимодействие РНКII с РНКI стабилизирует конформацию РНКII, не способную связываться с ДНК

причем РНКІ транскрибируется в противоположном направлении, т. е. РНКІІ и РНКІ комплементарны друг другу (рис. 38, а). Комплементарное спаривание РНКІ с РНКІІ так меняет вторичную структуру последней, что она не может гибридизоваться с областью инициации синтеза ДНК. Интересно, что участок регуляторного воздействия, спаривания РНКІІ с РНКІ отстоит на несколько сотен нуклеотидов от того места, где происходит регулируемое событие — гибридизация РНК с ДНК. Как и в других аналогичных случаях (например, при регуляции активности генов биосинтеза аминокислот, см. гл. VII), дело здесь в том, что РНКІІ может в принципе образовать два различных, причем взаимоисключающих варианта вторичной структуры. Лишь один из вариантов способен гибридизоваться с ориджином, а спаривание с РНКІ стабилизирует другой вариант и тем самым исключает репликацию. Вторичная структура РНК образуется по мере ее синтеза РНК-полимеразой, т. е. начиная с 5'-конца, поэтому событие, происходящее в 5'-концевой области РНКІІ, — спаривание с РНКІ — существенно влияет на вторичную структуру удаленных областей РНКІІ (рис. 38, б).

Взаимодействие РНКІ и РНКІІ — процесс сложный. Их непосредственному спариванию мешает развитая вторичная структура обеих РНК. Сначала взаимодействуют одноцепочечные петли на концах шпилек, причем последовательность этих петель определяет скорость взаимодействия и, следовательно, степень подавления репликации. Затем, видимо, происходит изомеризация комплекса и полное комплементарное спаривание РНКІІ с РНКІ (рис. 38, в).

Регуляция репликации основана на том, что РНКІ, ингибитор репликации, нестабильна. Это значит, что количество РНКІ в клетке пропорционально скорости ее синтеза, т. е. количеству генов этой РНК или, другими словами, количеству копий плазмиды, а *концентрация* РНКІ, которая и определяет скорость взаимодействия с РНКІІ и частоту инициации, должна быть пропорциональна количеству копий плазмиды и обратно пропорциональна объему клетки. Когда клетки растут и увеличивается их объем, концентрация РНКІ падает и становятся возможными новые раунды инициации. После того как произошла репликация, количество копий плазмиды возрастает и одновременно возрастает концентрация РНКІ, что подавляет новые инициации до тех пор, пока рост клеток снова не приведет к достаточному разбавлению РНК-ингибитора.

Описанная система обратной связи поддерживает определенную копийность репликона в клетке. В принципе репликация может контролироваться дополнительными факторами, изменяющими число копий в зависимости от условий. Так, ген *rop ColE1* кодирует белок-репрессор, подавляющий синтез РНКІІ и, по-видимому, усиливающий взаимодействие РНКІІ с РНКІ. Очевидно, что действие репрессора направлено на уменьшение числа копий плазмиды. Физиологический смысл этой регуляторной системы пока не ясен.

Контроль копийности с помощью РНК-ингибитора, влияющей на вторичную структуру другой РНК, необходимой для инициации, характерен не только для плазмид типа ColE1. Другие варианты той же стратегии используют, например, плазмиды *E. coli* R1 и стафилококковая плаزمиды рТ181.

Многие репликоны используют, по-видимому, совершенно иную стратегию регуляции собственного синтеза. Для инициации репликации этих репликонов необходим белок-инициатор (например, белок F в случае плазмиды F). Этот белок специфически связывается с определенной последовательностью ДНК, многократно повторенной на данном репликоне. Связывание белка-инициатора с одной или несколькими такими последовательностями, находящимися в ориджине, необходимо для инициации. Одна из последовательностей находится в начале гена белка-инициатора, так что связывание с ней белка подавляет его собственный синтез. Считается, что регуляция репликации осуществляется благодаря сложной конкуренции за белок-инициатор между участком ДНК, необходимым для собственной репрессии, участком (или участками), необходимым для инициации синтеза ДНК, и другими участками связывания. Хотя подобные репликоны пока еще недостаточно изучены и детальная картина регуляции репликации не ясна, очевидно, что наличие множественных мест связывания ключевого белка инициации репликации позволяет регуляторной системе очень чутко отзываться на изменение копийности репликона. Например, если плазмиды содержит 10 повторенных мест связывания белка-инициатора, то появление за счет репликации одной дополнительной копии плазмиды увеличит число участков связывания на 10. В определенном смысле многократно повторенные участки связывания белка-инициатора, суммарное количество которых пропорционально копийности репликона, аналогичны ранее рассмотренной ингибиторной РНК, концентрация которой также пропорциональна копийности.

Не исключено, что описанную стратегию регуляции репликации использует и хромосома *E. coli*. В хромосомном ориджине, *oriC*, имеется несколько участков связывания белка-инициатора *dnaA*. Кроме того, известно, что *dnaA*-белок является репрессором собственного синтеза.

Бактериальная хромосома отличается от плазмидных репликонов тем, что ввиду ее больших размеров раунд репликации занимает промежуток времени, по продолжительности сравнимый с

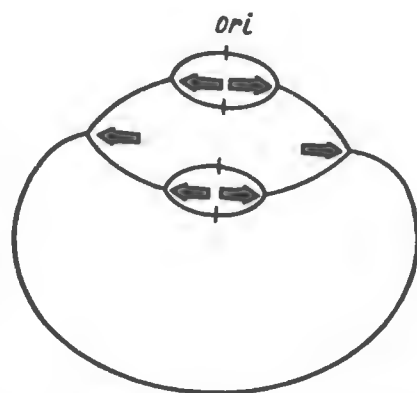


Рис. 39. Структура реплицирующейся бактериальной хромосомы

(При высокой скорости роста бактерий инициация нового раунда репликации происходит еще до окончания предыдущего. Такая «дихотомическая» репликация позволяет бактериям при благоприятных условиях иметь время генерации меньшее, чем время, необходимое на завершение полного раунда репликации ДНК)

временем генерации. Более того, при максимальной скорости роста время удвоения бактерий меньше времени полной репликации хромосомы! Это оказывается возможным благодаря тому, что на хромосомном ориджине инициация может произойти еще до завершения предыдущего раунда репликации. При малой скорости роста, когда инициация происходит редко, ДНК бактерий полностью реплицируется между двумя последовательными актами инициации, а при высокой скорости роста инициации происходят часто, еще до того, как закончится предыдущий раунд репликации. В результате возникает разветвленная структура, показанная на рис. 39. При такой *дихотомической* репликации возрастает количество одновременно работающих репликативных вилок и пропорционально возрастает скорость синтеза ДНК.

6. СЕГРЕГАЦИЯ БАКТЕРИАЛЬНЫХ РЕПЛИКОНОВ ПО ДОЧЕРНИМ КЛЕТКАМ ПРИ ДЕЛЕНИИ

Стабильное поддержание любого репликона требует не только согласования его репликации с клеточным делением, но и упорядоченного распределения молекул ДНК по дочерним клеткам. Считается, что правильная сегрегация достигается у бактерий за счет прикрепления ДНК к мембране, причем пространственная организация

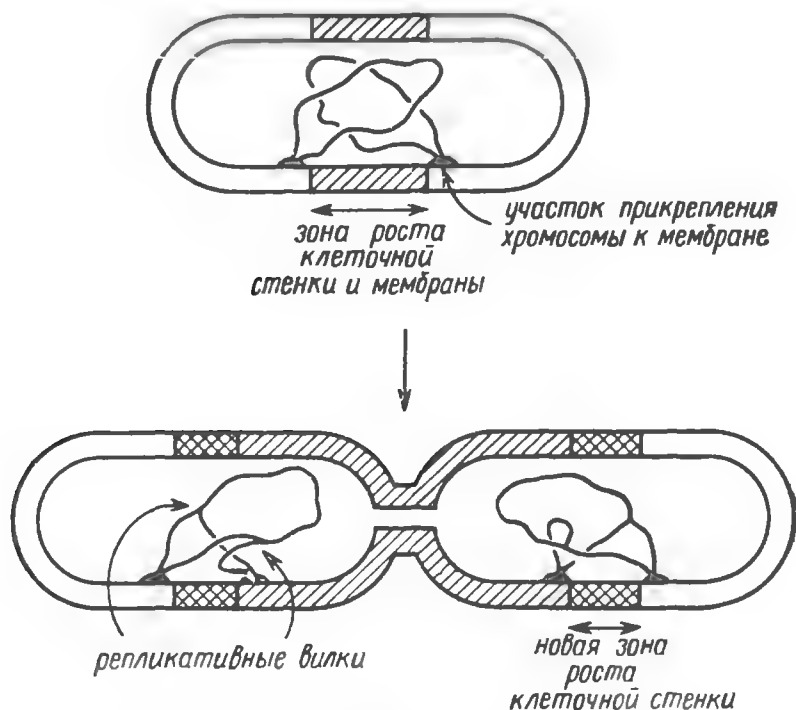


Рис. 40. Механизм упорядоченной сегрегации
(Упорядоченная сегрегация дочерних молекул ДНК по клеткам при делении предположительно достигается за счет прикрепления ДНК к мембране)

участков прикрепления и зоны роста мембраны и клеточной стенки обеспечивает автоматическое «растаскивание» двух копий прореплицировавшейся молекулы ДНК по разным дочерним клеткам (рис. 40). Подобная стратегия, видимо, характерна не только хромосомам, но и низкокопийным плазмидам, например F-фактору *E. coli*.

Многие мелкие репликоны используют альтернативную стратегию стабильного наследования. Они, по-видимому, не имеют механизма упорядоченной сегрегации, но поддерживаются в высоком числе копий. Высокая копияность обеспечивает относительно стабильное наследование репликона при случайном распределении молекул по дочерним клеткам при делении. Вероятность того, что репликон при таком способе распределения будет утерян (т. е. в одну из дочерних клеток не попадет ни одной копии ДНК данного репликона), равна $(1/2)^n$, где n — число копий, т. е. меньше 0,1 % уже для 10 копий плазмиды в клетке. Естественно, для этого способа принципиально важным является строгое восстановление копияности, чтобы единственная молекула ДНК успела быстро размножиться до характерного для нее числа копий до начала следующего деления клетки.

7. ОСОБЕННОСТИ РЕПЛИКАЦИИ ДНК ЭУКАРИОТ

Инициация репликации у эукариот

Эукариотические геномы, как правило, значительно больше бактериальных, а скорость синтеза ДНК у эукариот, напротив, существенно ниже (несколько десятков нуклеотидов в секунду). Видимо, по этой причине инициация синтеза ДНК эукариот происходит во многих разных точках хромосомы, т. е. эукариотические хромосомы имеют *полирепликонную организацию*. Характерный размер репликона высших эукариот — около 100 000 п. о. По крайней мере у низших эукариот, например у дрожжей, инициация происходит не на случайных последовательностях, а на вполне определенных ориджинах репликации. Эти участки были в ряде случаев охарактеризованы (у дрожжей они называются ARS — от англ. *autonomous replicating sequences* — автономно реплицирующиеся последовательности), они различны между собой по последовательности, но имеют несколько участков гомологии, видимо, служащих местами узнавания ответственных за инициацию репликации ядерных белков. Происходит ли инициация у высших эукариот на вполне определенных ориджинах, не ясно, но у многих эукариотических вирусов дело обстоит именно так (см. гл. XIII).

Инициация репликации строго регулируется. Полирепликонная организация требует, чтобы в каждом цикле клеточного деления каждый ориджин «сработал» только один раз, в противном случае на хромосоме образуются разветвленные структуры. Для дрожжевых

ARS-последовательностей показано, что инициация действительно происходит один раз в каждой S-фазе клеточного цикла. У высших эукариот достаточно протяженные участки хромосом, содержащие несколько соседних репликонов, начинают синтез ДНК в S-фазе приблизительно одновременно, но каждый такой участок имеет характерное для него время инициации репликации: одни активируются в начале S-фазы, другие — позже, а третьи — в конце. В этом случае особенно ясно видно, что существует запрет на повторную инициацию репликации в том же цикле клеточного деления, хотя одновременно может происходить инициация на других, не сработавших ранее ориджинах. Некоторые данные указывают на то, что подобный запрет снимается в ходе митоза. Не исключено, что белки, лимитирующие инициацию, не имеют доступа в ядро и способны связаться с ДНК предполагаемых ориджинов лишь в ходе митоза, когда ядерная оболочка распадается. Нарушение запрета на

повторную инициацию, которое очень редко, но все же происходит, приводит к амплификации (появлению множественных копий) того или иного участка хромосомы, где произошла «лишняя» инициация репликации.

Репликация концов эукариотических хромосом

Неспособность ДНК-полимераз иницировать новые цепи и то, что они умеют удлинять только 3'-конец, делают невозможной полную репликацию линейной ДНК (рис. 41). В каждом раунде репликации линейные молекулы должны укорачиваться. Хромосомы эукариот линейны, поэтому должен существовать специальный механизм репликации их концов — *теломер*. Уже давно генетики обратили внимание на то, что теломеры хромосом стабильны в отличие от концов двуцепочечных разрывов ДНК, которые отличаются нестабильностью и вызывают рекомбинации. На

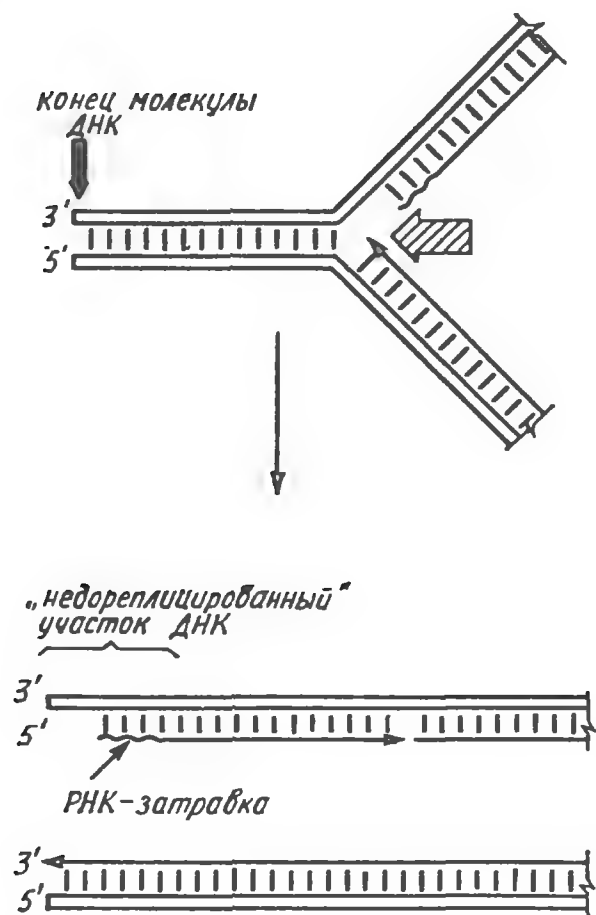


Рис. 41. Схема, демонстрирующая сложности, возникающие при репликации линейных молекул ДНК

(Репликация запаздывающей цепи не может пройти полностью, так как даже если РНК-затравка и может образоваться на самом конце матричной цепи, то после ее удаления ни одна ДНК-полимераза не сможет восстановить незастроенный участок)

этом основании был сделан вывод об особом строении теломер хромосом.

Теломерные области были выделены из ДНК ряда низших эукариот — инфузорий и дрожжей. Оказалось, что на концах хромосом этих организмов имеются простые повторяющиеся последовательности, например у тетрахимены теломерная последовательность такова: $(TTGGGG)_n$ [соответственно $(CCCCAA)_n$ в комплементарной цепи ДНК]. Эта повторенная последовательность «загнута» сама на себя за счет неканонических G·G-пар оснований и образует не совершенную шпильку, так что свободных 5'- и 3'-концов у хромосомы нет. Имеющийся в ядре фермент безматрично присоединяет к 3'-концу хромосомной ДНК последовательность концевой повтора, причем фермент работает лишь в том случае, если на конце ДНК теломерные повторы уже присутствуют (рис. 42). Поэтому, вместо того чтобы укорачиваться в каждом раунде репликации, хромосомы на самом деле удлиняются. Каким образом контролируется количество простых повторов на конце хромосомы, не ясно. Ясно, од-

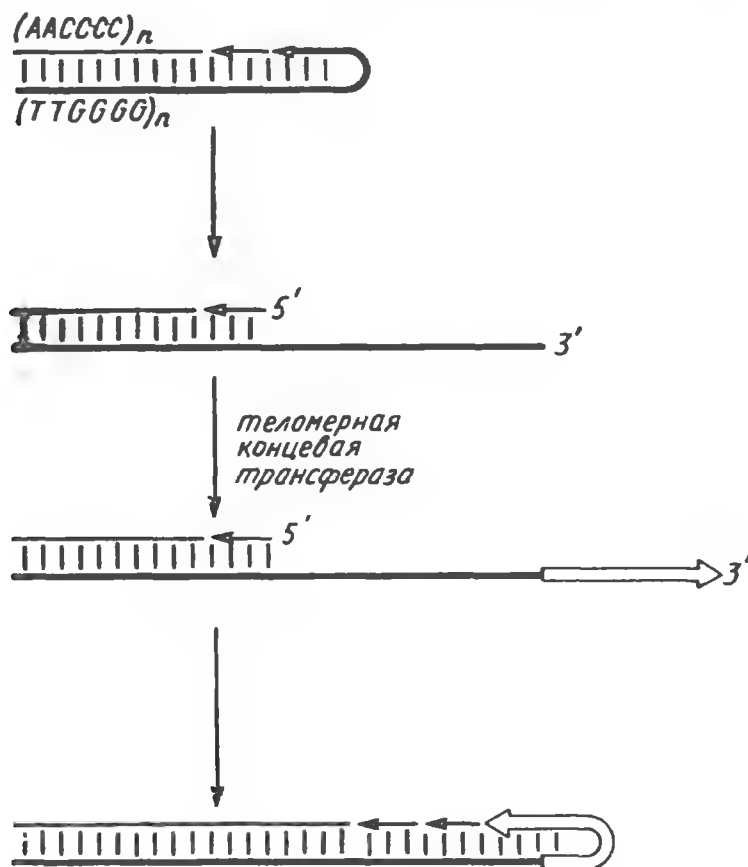


Рис. 42. Модель репликации теломер (на примере тетрахимены)

(Цепь, состоящая из повторенных последовательностей 5'TTGGGG3', замкнута сама на себя за счет образования неканонических G·G-пар. После того как она разворачивается, ее удлиняет терминальная трансфераза. Удлинившаяся матрица позволяет праймазе и полимеразе удлинить комплементарную цепь; оставшийся одноцепочечным 3'-конец вновь загибается сам на себя)

нако, что должен существовать механизм, не дающий хромосоме расти до бесконечности. Показано, что у трипаносом концы хромосом удлиняются при каждой раунде репликации, но время от времени происходит их существенное укорочение. При размножении дрожжей хромосомы не растут, но в одном из дрожжевых мутантов с нарушением нормального клеточного цикла на непермиссивной температуре начинается удлинение теломер. Можно думать, что ген, маркированный этой мутацией, участвует в ограничении длины хромосом.

Данные последнего времени указывают на аналогичную организацию теломер у высших животных и растений.

Сегрегация эукариотических хромосом при делении клеток

Известно, что в мейозе и в митозе хромосомы упорядоченно расходятся по дочерним клеткам с помощью аппарата веретена, микротрубочки которого обеспечивают «растягивание» дочерних хромосом или гомологов к разным полюсам. Микротрубочки веретена прикрепляются к специальному участку хромосомы — *кинетохору*. Это белковый комплекс, который собирается на специализированной последовательности хромосомной ДНК — *центромере*. Молекулярные основы функционирования кинетохора пока не ясны. Методы молекулярного клонирования позволили выделить центромеры хромосом дрожжей. Вставление этих последовательностей в способные реплицироваться молекулы ДНК обеспечивает правильную сегрегацию последних в митозе у дрожжей. В случае дрожжей-сахаромицетов центромеры оказались сравнительно короткими (100—200 п. н.) сегментами ДНК. Центромеры делящихся дрожжей значительно больше (несколько тысяч п. н.) и, видимо, напоминают своим строением центромеры высших эукариот. Механизм упорядоченной сегрегации хромосом эукариот станет понятен, когда выяснится, как связанные с центромерой кинетохорные белки взаимодействуют с аппаратом веретена.

Поскольку с помощью методов клонирования у дрожжей были выделены все элементы, необходимые для репликации и наследования хромосом — *ориджины репликации*, *теломеры* и *центромеры*, — то оказалось возможным создать *искусственную хромосому*, состоящую из соединенных вместе двух теломер, центромеры, последовательности *ARS* и ДНК «наполнителя», роль которой может играть ДНК с любой последовательностью, например ДНК фага лямбда. Оказалось, что искусственная хромосома поддерживается в дрожжах, причем стабильность ее наследования не намного ниже, чем стабильность собственных дрожжевых хромосом.

ЛИТЕРАТУРА

А. Корнберг. Синтез ДНК.— М.: Мир, 1977.

Б. Альбертс, Д. Брей, Дж. Льюис и др. Молекулярная биология клетки.— М.: Мир, 1987. Т. 2.

Б. Льюин. Гены.— М.: Мир, 1987.

Глава III

РЕПАРАЦИЯ ДНК

Видимо, уже на ранних стадиях эволюции ДНК заменила РНК в качестве носителя генетической информации. Этому гипотетическому событию должны были способствовать большая химическая устойчивость ДНК, связанная с заменой рибозы на дезоксирибозу, и двуцепочечное строение, «скрывающее» целый ряд реакционноспособных группировок. Но несмотря на свои «преимущества», ДНК постоянно подвергается химическим изменениям, как спонтанным, так и индуцируемым мутагенами и даже клеточными метаболитами. Еще одна обычная причина повреждений ДНК — радиация и ультрафиолетовое облучение. Большинство происходящих с ДНК изменений недопустимы: они либо приводят к вредным мутациям, либо блокируют репликацию ДНК и вызывают гибель клеток. Поэтому все клетки имеют специальные системы исправления повреждений, репарации ДНК. Нарушение этих систем губительно. Репарация ультрафиолетовых повреждений ДНК нарушена у людей, страдающих тяжелым наследственным заболеванием — пигментной ксеродермой. Такие больные не могут бывать на солнце и обычно умирают в раннем возрасте от какого-либо злокачественного заболевания.

Принципы репарации ДНК у различных организмов сходны, поэтому рассмотрим эти принципы на примере *E. coli*, где они сравнительно хорошо изучены.

1. НАРУШЕНИЯ, ВОЗНИКАЮЩИЕ В ДНК

В ДНК сравнительно часто и спонтанно происходят *апуринизация* и дезаминирование оснований. Масштабы апуринизации можно оценить по тому, что ДНК каждой клетки человеческого организма ежедневно теряет около 5000 пуринов. Результатом апуринизации является АР-сайт (англ. apurinic-apurimidinic) — дезоксирибоза, лишенная основания.

При *дезаминировании* цитозин превращается в урацил, аденин — в гипоксантин, а гуанин — в ксантин. Наиболее вредно дезаминирование цитозина и аденина: обе реакции после репликации приводят к мутациям. Чаще всего дезаминируется цитозин; в ДНК каждой человеческой клетки за день происходит около 100 таких событий. При дезаминировании всех оснований возникают основа-

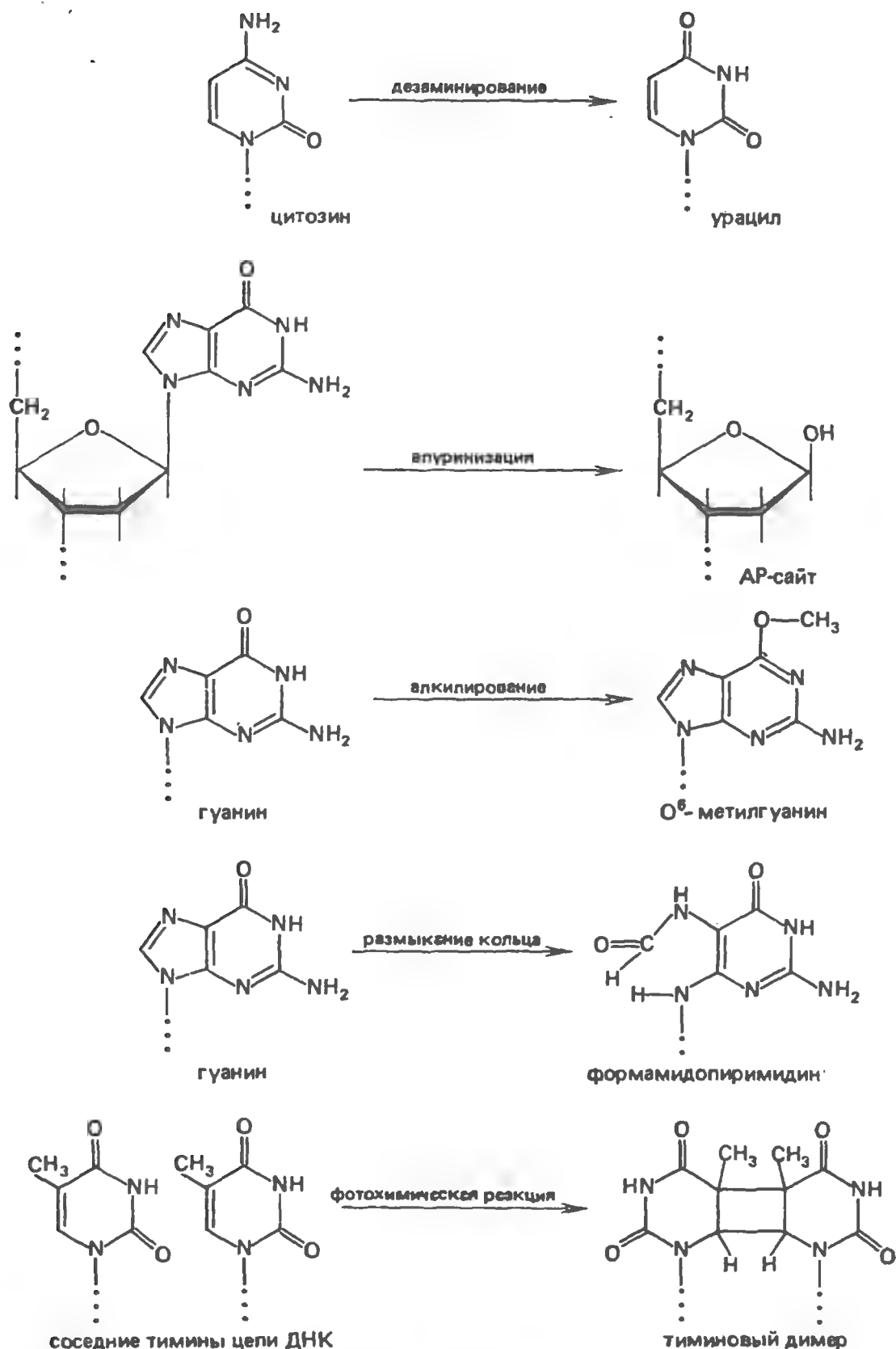


Рис. 43. Некоторые типы повреждений ДНК

(Показаны сравнительно часто возникающие повреждения. Апуринизация и дезаминирование происходят спонтанно. Алкилирование оснований — результат действия многих мутагенов. Пиримидиновые димеры образуются под действием ультрафиолетового облучения)

ния, не характерные для ДНК. Это обстоятельство позволяет репаративной системе клетки узнавать продукт дезаминирования и удалять его. Можно утверждать, что именно поэтому в ДНК в отличие от РНК вместо урацила присутствует тимин: урацил неотличим от продукта спонтанного дезаминирования цитозина.

Многие канцерогены *алкилируют*, например метилируют, основания ДНК. Наиболее частые продукты этих реакций — это O⁶-метилгуанин, 7-метилгуанин и 3-метиладенин. Первое из этих изменений мутагенно, а два других делают еще более лабильной N-гликозидную связь между основанием и сахаром, т. е. способствуют апуринизации.

Иногда может происходить *размыкание пуринового кольца*. Продукт формамидопиримидин представляет затруднение для репликации.

Главным нарушением, возникающим под действием ультрафиолета, является насыщение двойных связей оснований и образование *пиримидиновых димеров* из двух соседних пиримидинов цепи ДНК.

Различные типы вредных химических повреждений оснований ДНК показаны на рис. 43.

2. ПРЯМАЯ РЕАКТИВАЦИЯ ПОВРЕЖДЕНИЙ ДНК

Ряд повреждений клетка удаляет из ДНК путем *прямой реактивации*. Так, по крайней мере у бактерий существует специальный фермент, *метилтрансфераза*, который переносит метильную или этильную группу с алкилированного основания на один из собственных цистеиновых остатков. Алкилированный в результате собственной активности белок инактивируется, но может служить регулятором активности своего гена и нескольких других (см. раздел 4 этой главы).

У многих организмов обнаружен фотореактивирующий фермент (PRE, англ. photoreactivating enzyme), или *фотолиаза*, репарирующий пиримидиновые димеры. Фотолиаза *E. coli* — это полипептид с молекулярной массой 35 кД, прочно связанный с маленькой молекулой РНК (10—15 нуклеотидов), необходимой для активности фермента. Фермент находит на ДНК образовавшийся под действием ультрафиолетового облучения (220—320 нм) пиримидиновый димер и прочно связывается с ним. При облучении такого комплекса видимым светом (300—600 нм) происходит фотохимическая реакция «расщивания» димера, обратная реакция его образования. После того как нормальная структура ДНК восстановлена, фермент теряет к ней сродство.

Другие виды репарации ультрафиолетовых повреждений ДНК называют *темновой репарацией*, чтобы отличить от прямой фотоактивации.

3. ЭКСЦИЗИОННАЯ РЕПАРАЦИЯ

Если невозможна прямая реактивация, работают механизмы *эксцизионной репарации*, удаляющие из ДНК нарушенные участки. Так, практически на каждое аномальное основание, которое может возникнуть в ДНК, существует своя ДНК-N-гликозилаза — фермент, с высокой специфичностью узнающий в ДНК определенное аномальное основание, например формамидопиримидин, и вырезающий его с разрывом N-гликозидной связи (рис. 44). У *E. coli* есть около 20 различных ДНК-N-гликозилаз. Роль этих ферментов в репарации подтверждена исследованием мутантов: мутанты *E. coli* по гену *ung*, кодирующему урацил-ДНК-гликозилазу, характеризуются повышенным уровнем спонтанного мутагенеза; повреждение гена *tag*, кодирующего 3-метиладенин-ДНК-гликозилазу, повышают чувствительность клеток к мутагену метилметансульфонату.

После действия ДНК-N-гликозилазы в ДНК остается АР-сайт. Репарация такого повреждения может идти одним из двух путей. У эукариот был обнаружен фермент ДНК-инсертаза, способный непосредственно пришивать к дезоксирибозе основание в соответствии с комплементарной цепью ДНК. Пока обнаружены лишь пуриновые инсертазы, видимо вследствие того, что апуринизация происходит значительно чаще, чем потеря пиримидинов. Вторым путем репарации открывают ферменты, обнаруженные у всех организмов, — АР-эндонуклеазы. АР-эндонуклеазы разрезают саха-

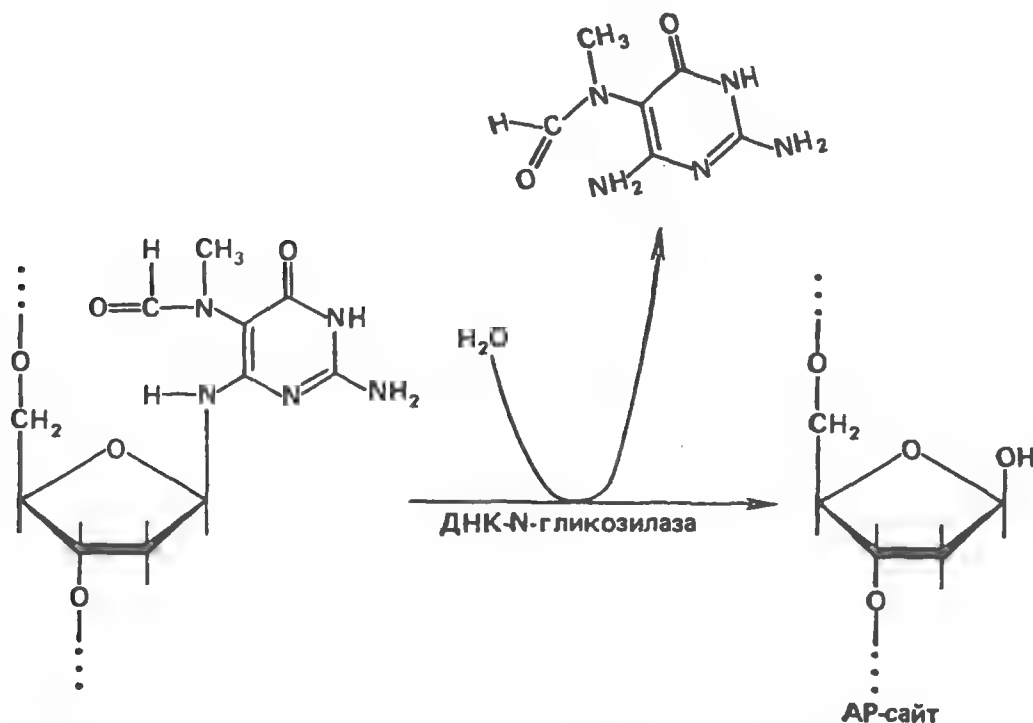


Рис. 44. Реакция, катализируемая ДНК-N-гликозилазами
(Существует множество ДНК-N-гликозилаз, специфичных к разным аномальным основаниям)

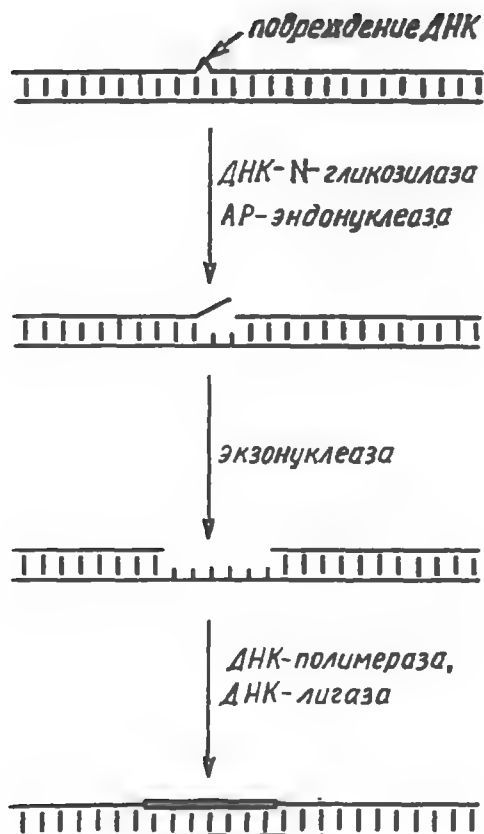


Рис. 45. Эксцизионная репарация (Поврежденный участок ДНК вырезается и застраивается заново с использованием комплементарной цепи в качестве матрицы. У *E. coli* ДНК-полимеразой, участвующей в репарации, служит ДНК-полимераза I. В ряде случаев этот же фермент может выполнять функции экзонуклеазы за счет своей 5' → 3'-экзонуклеазной активности)

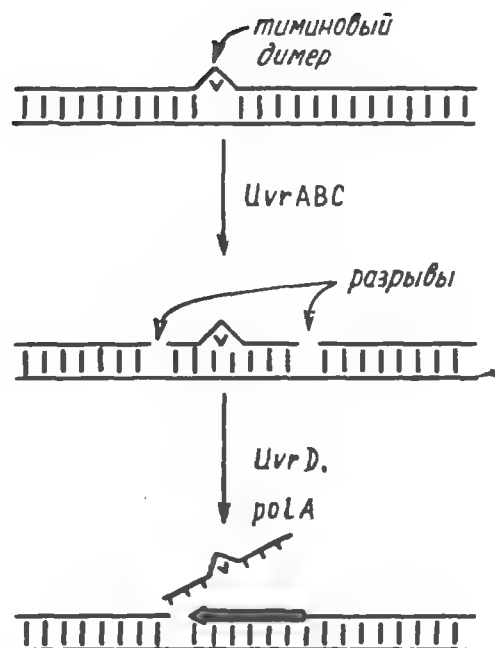


Рис. 46. Репарация повреждений, заметно нарушающих вторичную структуру ДНК

(АТФ-зависимая эндонуклеаза *UvrABC* вносит разрывы по обе стороны от повреждения. Геликаза *UvrD* способствует освобождению вырезанного олигонуклеотида; одновременно ДНК-полимераза I застраивает образовавшуюся брешь)

рофосфатный остов ДНК в AP-сайт. Существует два типа этих ферментов — одни рвут фосфодиэфирную связь с 3'-конца от AP-сайта, другие — с 5'-конца. В некоторых случаях AP-эндонуклеаза и ДНК-N-гликозилаза совмещены на одном полипептиде. Например, эндонуклеаза III *E. coli* — это AP-эндонуклеаза и гликозилаза, специфичная к дигидро- и дигидрокситимину — продуктам ультрафиолетового облучения. Разорванная после действия AP-эндонуклеазы цепь ДНК подвергается действию экзонуклеазы, удаляющей поврежденный участок. По крайней мере в некоторых случаях эту реакцию, видимо, осуществляет 5'-экзонуклеазная активность ДНК-полимеразы I, а полимеризующая активность заполняет образующуюся брешь. В том случае, если работают другие экзонуклеазы, брешь заполняет тоже ДНК-полимераза I. Остающийся разрыв сшивает ДНК-лигаза (рис. 45). Мутанты *E. coli* с нарушенными ДНК-полимеразой I или ДНК-лигазой дефектны по эксцизионной репарации.

Несколько отличный путь используется для репарации повреждений ДНК, заметно нарушающих структуру молекулы, например пиримидиновых димеров, образующихся под действием ультрафиолета. Такие повреждения удаляет специальный фермент — эндонуклеаза UvrABC (в темноте, когда не работает фотолиаза или когда повреждений в ДНК очень много). Эта нуклеаза разрывает фосфодиэфирные связи с 5'- и с 3'-конца от поврежденного участка, а затем с помощью белка UvrD, хеликазы II, поврежденный участок удаляется сопряженно с гидролизом АТР. Образующуюся брешь застраивает ДНК-полимераза I (рис. 46).

Видно, что эксцизионная репарация всегда использует один принцип: нарушенный участок ДНК удаляется, а затем восстанавливается на матрице ненарушенной комплементарной цепи ДНК. Другими словами, эксцизионная репарация основана на наличии в ДНК двух взаимокomплементарных цепей; более того, необходимость репарации — одна из причин, обусловивших наличие в ДНК двойного избытка информации.

4. ИНДУЦИРУЕМАЯ РЕПАРАЦИЯ

В условиях, увеличивающих количество повреждений ДНК, происходит индукция дополнительных репаративных ресурсов клетки. Классический пример индукции репарации — так называемая *реактивация Уэйгла*. Это явление состоит в том, что облученный ультрафиолетом бактериофаг может размножаться только на тех бактериях, которые тоже облучены ультрафиолетом. Это значит, что для успешного развития фага необходима репарация его ДНК, которую могут осуществить лишь клетки с индуцированной системой репарации. Аналогичное явление наблюдается и для эукариотических клеток и их вирусов. У бактерий индуцируемая репарация используется лишь в тех случаях, когда повреждений в ДНК становится настолько много, что это начинает угрожать клетке гибелью. Поэтому индуцируемая система репарации называется SOS-системой или системой SOS-репарации (табл. 5).

Степень индукции SOS-системы определяется количеством повреждений в ДНК: при небольшом количестве повреждений возрастает уровень некоторых репаративных белков, работавших и до индукции, например UvrA, B, C и D. При большем количестве повреждений блокируется деление клеток (в норме оно восстанавливается, если клетке удалось «починить» ДНК) и индуцируется синтез белка-продукта гена *recA*, необходимого для рекомбинационной репарации и для дальнейшей индукции SOS-системы (см. ниже). При еще большем количестве повреждений ДНК индуцируются гены *umiC* и *umiD*, которые ответственны за особый путь репарации, сопряженный с возникновением мутаций. Механизм действия продуктов генов *umiCD* не ясен, но предполагается, что они позволяют

Таблица 5. Некоторые гены SOS-системы и результат их индукции

Индукцируемый ген	Физиологическое действие
<i>uvrA</i> , В, С и D	Репарация повреждений, нарушающих вторичную структуру ДНК, реактивация Уэйгла
<i>recA</i>	Пострепликативная репарация (см. гл. IV, раздел 1), репарация двуцепочечных разрывов, мутагенез, индукция SOS-системы
<i>recX</i>	Репарация двуцепочечных разрывов
<i>ruv</i>	То же
<i>ssb</i>	Участвует в репарации и рекомбинации
<i>umuDC</i>	Мутагенез, реактивация Уэйгла
<i>sulA</i>	Подавление клеточного деления (филаментация)
<i>lexA</i>	Выключение SOS-ответа
?	Ослабление рестрикции (см. гл. VI)
?	Прекращение дыхания

ДНК-полимеразе копировать нарушенный участок матрицы (рис. 47). Для этого продукты генов *umuCD* и *recA* должны изменить свойства ДНК-полимераз. Оказалось, что корректирующая активность ϵ -субъединицы ДНК-полимеразы III специфически подавляется RecA-белком, связанным с поврежденным участком матрицы, например тиминовым димером. Можно думать, что продукты генов *umuCD* позволяют ДНК-полимеразе «в виде исключения» удлинять затравку, последний нуклеотид которой не спарен с матрицей. Подобное совместное действие продуктов генов *umuCD* и RecA-белка может быть полезным для выживания в условиях, когда поврежде-

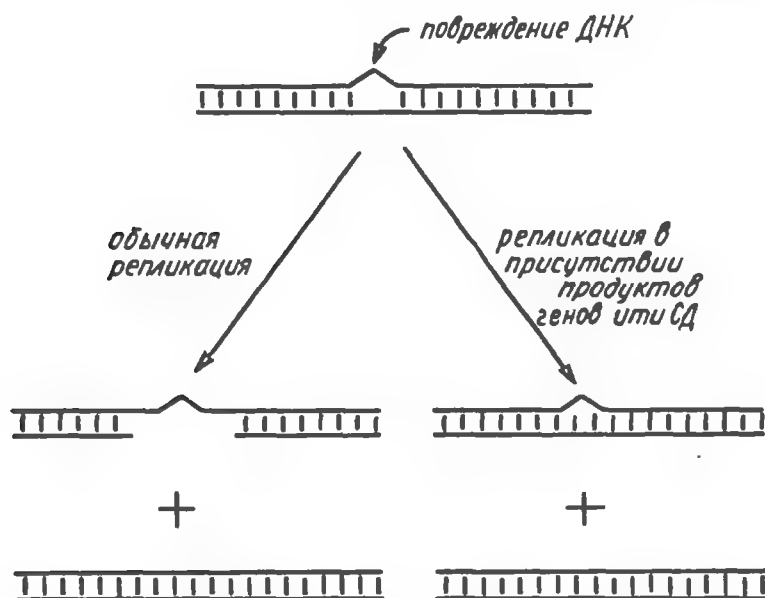


Рис. 47. Предположительное действие продуктов генов *umuDC*

(«Копирование» поврежденного участка матрицы сопровождается образованием мутаций)

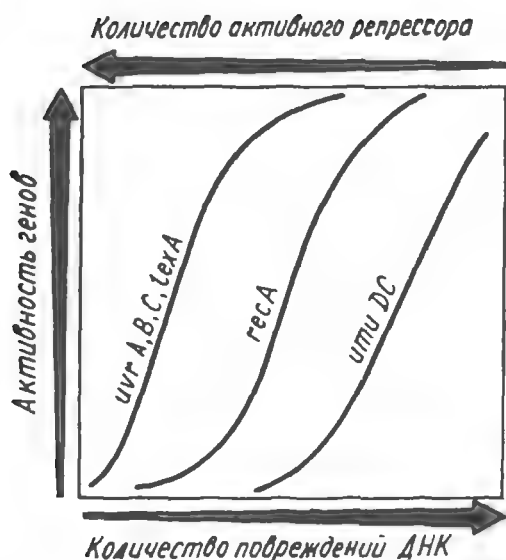


Рис. 48. Схематический график, иллюстрирующий тонкую настройку SOS-системы

(Степень индукции разных генов, входящих в SOS-систему, по-разному зависит от количества повреждений ДНК)

ний в ДНК очень много, в то же время очевидно, что активность генов *umuCD* приводит к появлению мутаций, поскольку при репликации напротив поврежденного участка матрицы, по-видимому, вставляется произвольный нуклеотид. В любом случае гены *umuD* и *C* обеспечивают весь мутагенный эффект ультрафиолетового облучения и многих химических мутагенов. Таким образом, мутагенез не является пассивным процессом, а требует активного вмешательства продуктов определенных клеточных генов. Эти генные продукты «обрабатывают» первичное нарушение в ДНК и лишь после этого возникает мутация.

Механизм индукции SOS-системы *E. coli* основан на взаимодей-

ствии продуктов двух генов: *recA* и *lexA*. Белок LexA является репрессором и подавляет активность не менее 17 различных генов (часть из них приведена в табл. 5). Среди прочих LexA-белок подавляет собственный синтез и синтез белка RecA. Белок RecA способен связываться с поврежденными участками ДНК (например, с пиримидиновыми димерами в составе ДНК) или с продуктами деградации поврежденной ДНК (олигонуклеотидами), при этом RecA-белок приобретает способность связываться с белком LexA. Такое связывание изменяет конформацию белка LexA и активирует его скрытую автопротеазную активность, что приводит к расщеплению LexA-репрессора пополам. В результате репрессор инактивируется и происходит индукция всех подавляемых им генов, в том числе его собственного гена *lexA*. Когда исчезает индуцирующий сигнал (поврежденная ДНК), RecA-белок перестает связываться с LexA-репрессором и происходит быстрое возрастание концентрации последнего, поскольку ген *lexA* полностью индуцирован. В результате все гены SOS-системы подавляются. Индукция SOS-системы тонко настроена: для подавления разных генов SOS-системы нужны различные концентрации репрессора. Так как количество репрессора зависит от количества активированного RecA-белка, которое, в свою очередь, зависит от количества повреждений в ДНК, то степень индукции SOS-системы будет связана с количеством повреждений (рис. 48): при незначительных повреждениях ДНК индуцируется часть генов SOS-системы и лишь при очень сильном повреждении ДНК индуцируются все гены. Поскольку репарация — процесс достаточно энергоемкий, индукция SOS-системы происходит толь-

ко при условии благоприятного энергетического баланса клетки. Это обеспечивается тем, что RecA-белку для активации и связывания с LexA-репрессором необходим не только индуцирующий сигнал (поврежденная ДНК или продукты ее деградации), но и АТФ.

Степень индукции SOS-системы в определенном смысле отражают «благополучие» клетки и ее шансы на выживание. Поэтому некоторые относительно автономные внутриклеточные генетические элементы, например умеренные бактериофаги, используют индукцию SOS-системы в качестве сигнала для размножения и уничтожения клетки-хозяина: безвредный до того участок хромосомы (профаг, см. гл. XIII), «почувствовав» слабость хозяина, начинает размножаться и уничтожает его, чтобы спастись самому. Для фага лямбда показано, что чувствительность к состоянию индукции SOS-системы объясняется тем, что репрессор фага устроен аналогично белку LexA и «самораскусывается», связавшись с активированным RecA-белком.

SOS-система — это не единственная индуцируемая система репарации у *E. coli*. Совершенно другой набор генов индуцируется в ответ на алкилирующие соединения. Как было сказано, часть повреждений, вызываемых в ДНК этими соединениями, репарируется за счет прямой реактивации метилтрансферазой, которая сама при этом инактивируется. Оказалось, что алкилированная метилтрансфераза служит активатором транскрипции и повышает активность ряда генов, в том числе собственного гена *ada* и гена *alkA*, кодирующего ДНК-N-гликозилазу, специфичную к алкилированным основаниям. Обработка клетки небольшими количествами алкилирующих соединений вызывает 30-кратное увеличение уровня метилтрансферазы и предотвращает гибель клеток под действием существенно больших доз алкилирующих мутагенов.

5. РЕПАРАЦИЯ НЕСПАРЕННЫХ НУКЛЕОТИДОВ

Системы репликации и репарации ДНК хоть и очень редко, но все же ошибаются. В результате в ДНК может включиться «неправильный», т. е. не комплементарный матрице, нуклеотид. Другой источник неспаренных нуклеотидов — гомологичная рекомбинация, в ходе которой образуются гетеродуплексы, состоящие из двух комплементарных цепей, исходно принадлежавших разным молекулам ДНК (см. гл. IV). Такие нарушения структуры ДНК репарируются, по крайней мере у бактерий и низших эукариот. Система репарации должна каким-то образом отличать друг от друга две цепи одной молекулы ДНК, чтобы решить, какой из двух неспаренных нуклеотидов «правильный», и заменить «неправильный» нуклеотид на нуклеотид, комплементарный «правильному».

В случае нарушений спаривания, возникающих в результате репликации или репарации, новосинтезированная цепь ДНК неко-

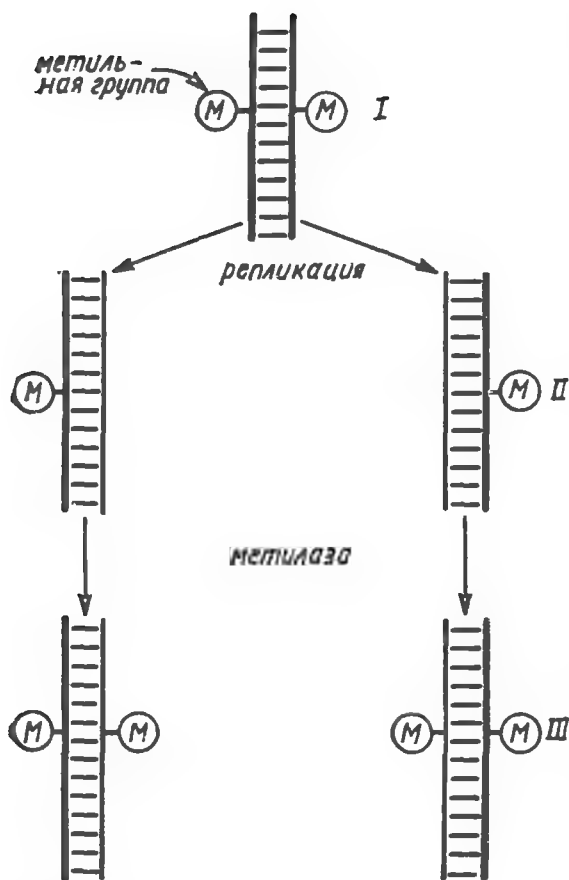


Рис. 49. Метилирование ДНК

(Метилирование позволяет отличить «старую» цепь ДНК от «новой». В течение нескольких минут после репликации, пока не действовала метилаза, новосинтезированная ДНК не метилировалась, т. е. отличается от матричной цепи): I — полностью метилированная ДНК; II — полуметилированная ДНК; III — полностью метилированная ДНК

торое время может отличаться от матричной цепи наличием одноцепочечных разрывов или брешей. Не исключено, что это различие использует система репарации, контролирующая правильность спаривания оснований. У *E. coli* за такой контроль ответственны продукты генов *recF* и *mutS*. Для их действия необходим АТР, что наводит на мысль о направленном движении комплекса репаративных белков по одной из цепей ДНК начиная с одноцепочечного разрыва.

Другое отличие между новосинтезированной и матричной цепями, по крайней мере у *E. coli*, обусловлено метилированием ДНК. Метилаза, кодируемая геном *dam*, метилирует аденины в последовательности GATC. Большинство таких последовательностей метилированы в обеих цепях ДНК. После репликации возникает полуметилированная ДНК: последовательности GATC метилированы только в матричной, но не в новой цепи (рис. 49). Со временем полумети-

лированные участки метилируются полностью, но пока отличие существует, оно служит указанием о направлении репарации неспаренных нуклеотидов для специализированной репарационной системы. В этой системе участвуют продукты генов *mutH*, *mutL* и *mutS*, SSB-белок и хеликаза II (UvrD-белок). Для работы системы необходим АТР и дезоксирибонуклеотиды. Скорее всего, *dam*-зависимая система репарации, обнаружив на одной молекуле ДНК неспаренные основания и полуметилированные участки, вносит в новосинтезированную ДНК разрыв и начиная от этого разрыва деградирует протяженный участок новой цепи, одновременно застраивая его заново (рис. 50). Мы еще раз убеждаемся в том, что поддержание высокой точности размножения и хранения информации обходится клетке дорого. Описанный механизм вносит вклад в эту точность. Мутации по генам *mutH*, *L* и *S* повышают частоту мутирования. К тому же эффекту приводят мутации по гену метилазы *dam*, или искусственно созданный в клетке избыток метилазы.

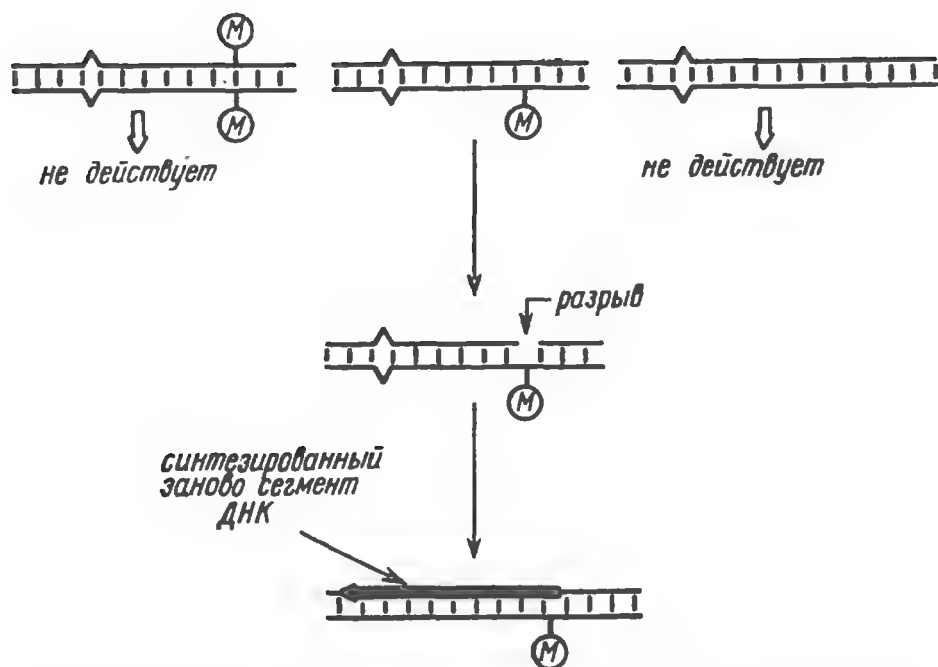


Рис. 50. Зависимая от метилирования система репарации неспаренных нуклеотидов

(Ферменты этого пути репарации действуют лишь на те молекулы, в составе которых есть и неспаренные нуклеотиды, и полуметилированные последовательности. Предположительно ферменты репарации вносят односторонний разрыв в неметилированную цепь полуметилированного участка, после чего деградируют протяженный сегмент разорванной цепи и одновременно застраивают его заново)

6. МЕТИЛИРОВАНИЕ ДНК И «ГОРЯЧИЕ ТОЧКИ МУТАГЕНЕЗА»

Метилирование не только способствует репарации ДНК, но может, напротив, облегчать мутагенез. В природе широко распространено метилирование цитозина. Каждый метилированный цитозин потенциально мутагенен, поскольку продуктом его спонтанного дезаминирования является тимин, нормальное для ДНК основание, а не урацил (продукт дезаминирования цитозина), который удаляется урацил-ДНК-гликозилазой (см. раздел 3 этой главы). У *E. coli* за счет активности гена *dcm* цитозины метилированы в последовательности $CC\text{A}^{\text{G}}\text{GG}$. Показано, что метилированный (внутренний) цитозин в этой последовательности действительно является «горячей точкой мутагенеза».

* * *

Дублирование информации в двух комплементарных цепях ДНК не позволяет безошибочно исправлять все типы повреждений. Описанные механизмы репарации не могут справиться с такими нарушениями структуры ДНК, как ковалентные межнитевые сшивки, которые могут возникать под действием ряда мутагенов, или двуцепо-

ческие разрывы ДНК, вызываемые, например, γ -облучением. Такие повреждения могут репарироваться только при наличии гомологичной неповрежденной молекулы ДНК, т. е. рекомбинационным путем. Рекомбинационная репарация будет рассмотрена в следующей главе.

ЛИТЕРАТУРА

Б. Альбертс, Д. Брей, Дж. Льюис и др. Молекулярная биология клетки.— М.: Мир, 1987. Т. 2.

Б. Льюин. Гены.— М.: Мир, 1987.

Л. Страйер. Биохимия.— М.: Мир, 1985. Т. 3.

Глава IV

РЕКОМБИНАЦИЯ

Рекомбинацией в общем смысле слова можно назвать возникновение новых последовательностей ДНК за счет разрывов и пересоединений предсуществующих молекул. Это понятие объединяет обширный круг разнообразных биологических явлений, роль которых будет рассмотрена по мере их описания.

1. ГОМОЛОГИЧНАЯ РЕКОМБИНАЦИЯ

Общая, или гомологичная, рекомбинация характерна для всех живых организмов от вирусов и бактерий до многоклеточных эукариот. При гомологичной рекомбинации происходит обмен участками между гомологичными, т. е. очень похожими по последовательности, молекулами ДНК. Так, к общей рекомбинации относятся обмены между гомологичными хромосомами в мейозе у эукариот и рекомбинационная инициация репликации ДНК бактериофага Т4 (см. гл. XIII). В первом приближении можно сказать, что гомологичная рекомбинация не создает принципиально новых последовательностей, а перетасовывает уже имевшиеся сходные варианты одной и той же последовательности (рис. 51). Чтобы подчеркнуть важность этого свойства, достаточно сказать, что при гомологичной рекомбинации между двумя сходными генами, кодирующими белок, оба рекомбинантных продукта оказываются не нарушенными, не происходит, например, сдвига рамки считывания. Другими словами, при гомологичной рекомбинации каким-то образом обеспечивается взаимное узнавание одинаковых (или очень сходных по последовательности) участков рекомбинирующих молекул. Если же гомологии нет, то и рекомбинация такого рода происходить не будет.

Как же обеспечивается взаимный поиск гомологии между рекомбинирующими молекулами ДНК? Наиболее естественное пред-

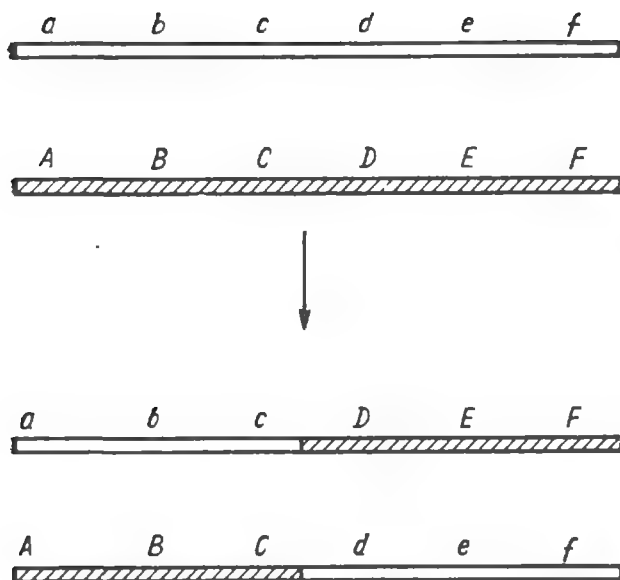


Рис. 51. Гомологичная рекомбинация

При рекомбинации такого рода происходит обмен частями между разными вариантами (обозначены большими и малыми буквами) одной последовательности. В результате возникают два новых рекомбинантных варианта той же последовательности

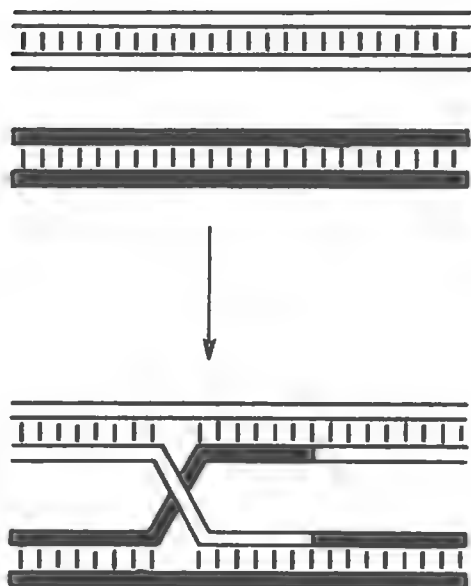


Рис. 52. Структура Холидея или полухиазма — промежуточное соединение реакции гомологичной рекомбинации

(Комплементарное спаривание цепей, принадлежащих разным дуплексам, обеспечивает взаимное узнавание гомологичных молекул)

положение состоит в том, что в этом процессе участвует комплементарное спаривание между цепями ДНК, принадлежащими разным рекомбинирующим дуплексам. Это предположение было достаточно четко экспериментально обосновано, и в настоящее время можно считать общепринятым, что гомологичная рекомбинация происходит через образование промежуточного соединения, называемого *структурой Холидея* или *полухиазмой* (рис. 52). В полухиазме осуществлено комплементарное спаривание между одноцепочечными участками, принадлежащими разным родительским молекулам ДНК. В результате в полухиазме имеются *гетеродуплексные районы*. Этот факт, как и многие другие особенности общей рекомбинации, был установлен в генетических опытах на дрожжах и других аскомицетах, поскольку при работе с этими организмами можно проследить за всеми продуктами индивидуального рекомбинационного события. Действительно, в результате мейоза у диплоидных дрожжей, гетерозиготных по какому-либо гену, образуются гаплоидные споры, которые иногда дают начало колонии, состоящей из двух секторов, одного — с одним аллелем родительской гетерозиготы и другого — с другим аллелем. Это значит, что спора, давшая начало такой колонии, несла оба аллеля в одной молекуле ДНК, т. е. соответствующий участок хромосомы этой споры был гетеродуплексным (рис. 53). Описанное явление называется *постмейотической сегрегацией*.

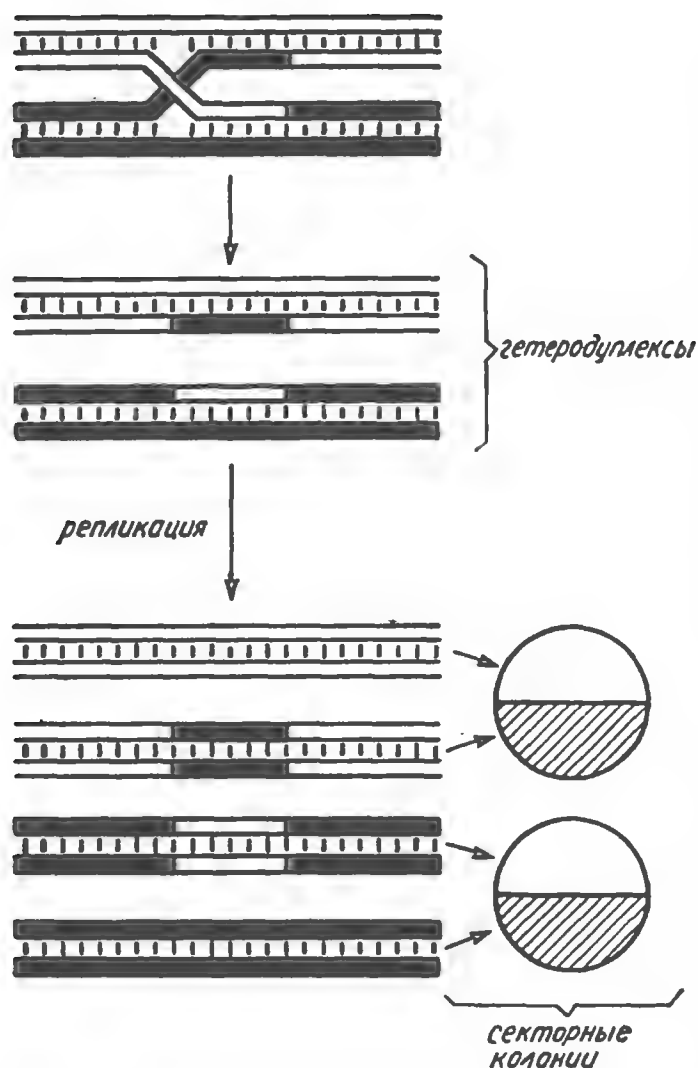


Рис. 53. Схематическое изображение постмейотической сегрегации

(Присутствующие в полухиазме гетеродуплексные районы остаются в молекулах ДНК — продуктах рекомбинации. Получившиеся в результате мейоза гаплоидные споры после первого деления дают две неодинаковые дочерние клетки: в одной последовательность ДНК соответствует одной цепи гетеродуплексного участка, во второй — соответствует второй цепи. Приведенная схема упрощена. Экспериментальные данные лучше объясняют модели, приведенные на рис. 64 и 67)

полухиазмы. Такой процесс может быть активным (см. ниже) или может происходить спонтанно, за счет вращательной диффузии; называется он *миграцией ветви* (от англ. branch migration; рис. 55).

Полухиазму можно разрезать двумя способами (рис. 56). В одном случае окажется, что рекомбинирующие молекулы обменялись сравнительно короткими одноцепочечными участками, образовавшими гетеродуплексные районы. При другом способе рекомбинирующие молекулы обмениваются своими частями, находящимися по обе стороны от места образования полухиазмы, часто именно этот случай называют *собственно рекомбинацией*.

Структуру Холлидея в качестве промежуточной стадии гомологичной рекомбинации удалось непосредственно наблюдать в электронный микроскоп. Действительно, если развернуть на 180° полухиазму, изображенную на рис. 52, то получится структура, показанная на рис. 54. Именно такие молекулы и удастся выявить при рекомбинации, причем объединены они по области гомологии.

Построение объемной молекулярной модели показало, что полухиазму можно образовать с сохранением всех водородных связей и стэкинг-взаимодействий между основаниями ДНК. Эта структура способна свободно перемещаться в обе стороны вдоль рекомбинирующих молекул ДНК. При этом обе гомологичные молекулы вращаются вокруг своей оси, а гетеродуплексный участок удлиняется или укорачивается в зависимости от направления смещения

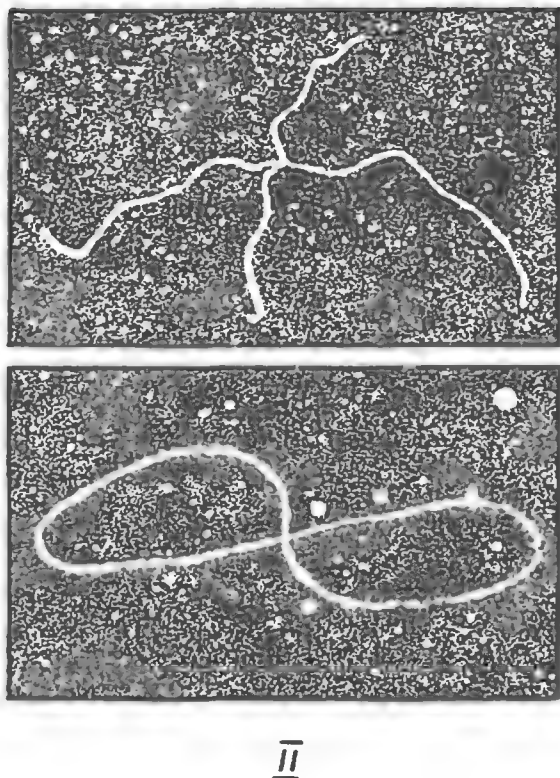
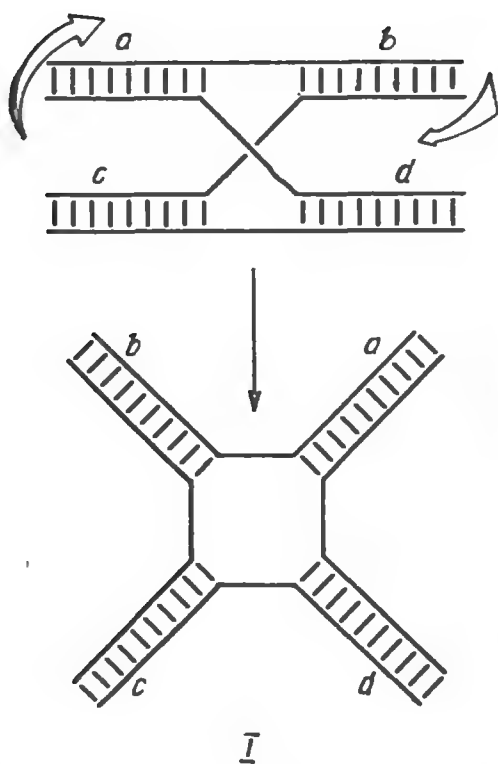


Рис. 54. Изомеризация полухиазмы

(Если развернуть полухиазму в направлении, указанном стрелками *I*, то получится структура, которую удастся непосредственно наблюдать в электронный микроскоп в качестве промежуточного рекомбинационного соединения *II*)

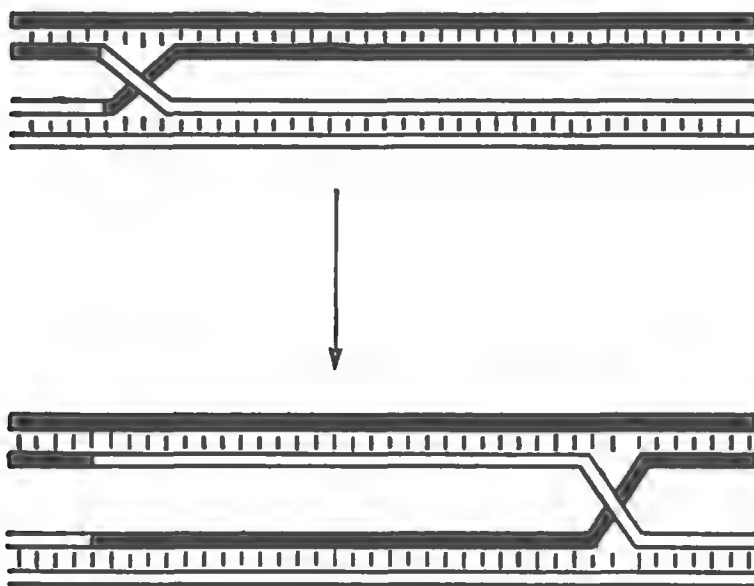


Рис. 55. Миграция ветви

За счет этого процесса может происходить удлинение (или укорочение) гетеродуплексного района. Так как реальные молекулы ДНК — это спирали, то миграция ветви сопровождается вращением обеих рекомбинирующих молекул

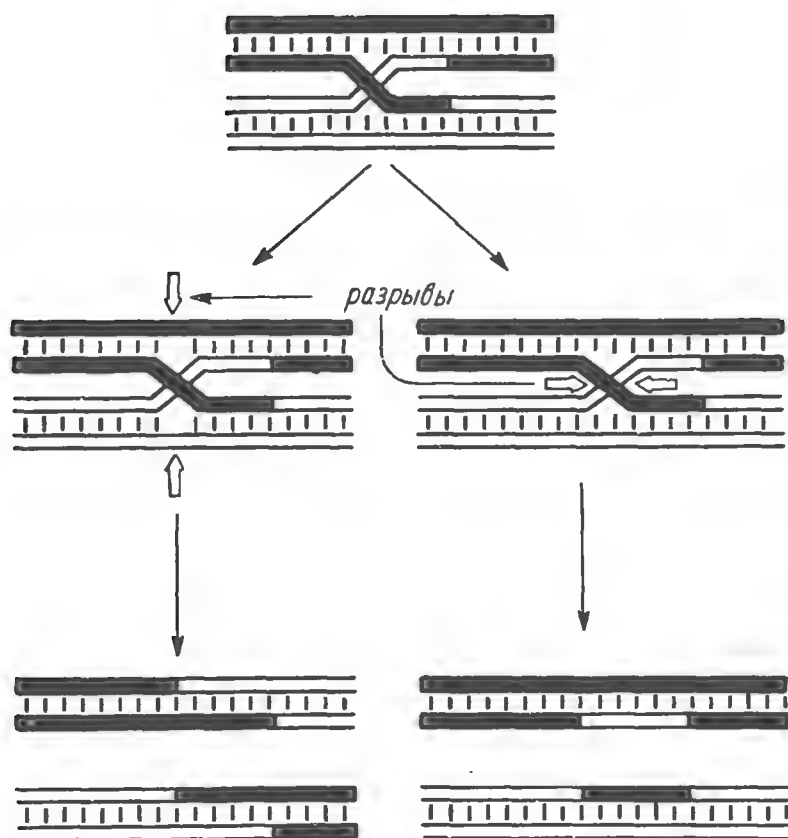


Рис. 56. Два способа разрезания полухиазмы

В зависимости от способа разрезания полухиазмы рекомбинирующие молекулы либо обмениваются сравнительно короткими одноцепочечными участками (справа), либо меняются частями, расположенными по обе стороны от места обмена одиночными цепями (слева)

Как уже говорилось, цепи ДНК в гетеродуплексе могут быть не полностью комплементарны друг другу. Судьба гетеродуплексных районов, образовавшихся в результате рекомбинации, может быть двоякой: если неспаренные основания гетеродуплекса не репарируются, то после репликации каждая цепь ДНК даст дуплекс с такой же последовательностью, как у молекулы ДНК, из которой она произошла (в случае рекомбинации в мейозе это приведет к постмейотической сегрегации); если же произойдет репарация неспаренных оснований, то информация, заключавшаяся в одной из цепей гетеродуплекса, исчезнет и заменится на информацию другой цепи. В отличие от других случаев рекомбинации в этом последнем случае один из двух возможных рекомбинантных продуктов исчезает и превращается (англ. convert) в другой, поэтому этот вариант рекомбинации называют *генной конверсией* (рис. 57).

Энзимология рекомбинаций у бактерий

Каким же образом происходит обмен комплементарными цепями между рекомбинирующими молекулами и образование холидеевской структуры? Этот вопрос был выяснен за последние несколько лет,

как и многие другие, на примере *E. coli*. У этого организма практически все виды гомологичной рекомбинации зависят от функционирования гена *recA* (этот ген играет важную роль и в репарации повреждений ДНК; см. гл. III). Оказалось, что RecA-белок (молекулярная масса около 40 кД) является ДНК-зависимой АТРазой, способной заменять одну из цепей дуплекса ДНК на гомологичную. Одна из реакций, которую может проводить очищенный RecA-белок *in vitro*, — образование так называемой D-петли (рис. 58). Для осуществления этой реакции RecA-белок покрывает одноцепочечную ДНК и такой комплекс связывается с двуцепочечной ДНК. Затем происходит перемещение однонитевой ДНК вдоль дуплекса до тех пор, пока не будет найден участок гомологии (если это возможно). Процесс поиска является активным и требует гидролиза АТР. Показано, что дуплекс ДНК в комплексе с RecA-белком и одноцепочечной ДНК заметно раскручен (вне зависимости от того, гомологична одонитевая ДНК дуплексу или нет); по всей вероятности, это облегчает поиск гомологии. После того как гомология найдена, RecA-белок проводит замену одной цепи ДНК на другую в активном процессе, требующем гидролиза АТР и имеющем вполне определенную направленность: перенос начинается с 3'-конца и продолжается в направлении 3'→5'. Другими словами, RecA-белок активно проводит одонитевую миграцию ветви. Более того, при наличии второго свободного конца RecA-белок способен обменивать цепи ДНК между двумя дуплексами, продвигая вдоль ДНК (миграция ветви) образовавшуюся таким образом структуру Холидея (рис. 59)!

Мы видим, что по крайней мере у бактерий в общих чертах понятен механизм центрального события гомологичной рекомбинации поиска гомологии и образования полухиазмы. Чтобы окончательно разобраться в молекулярных основах рекомбинации, необходимо рассмотреть, как происходит разрезание полухиазмы и как индуцируется рекомбинация — откуда берется свободный конец

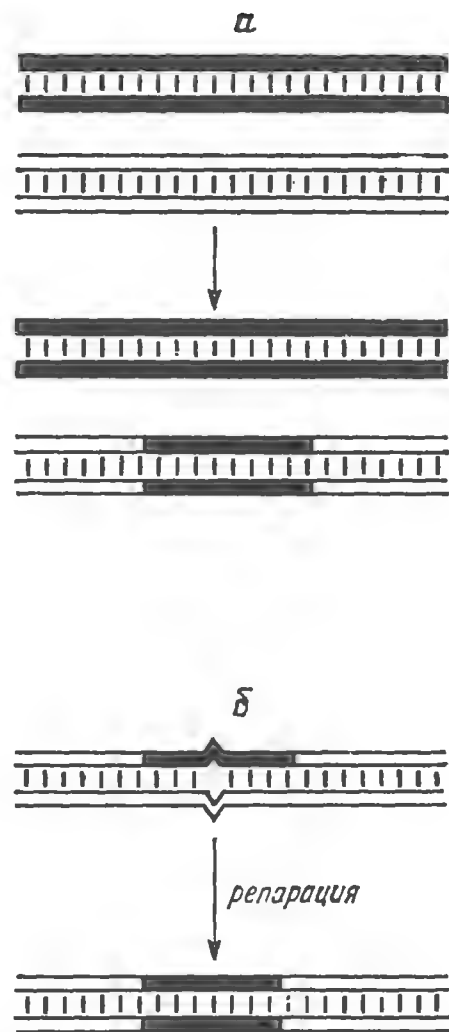


Рис. 57. Генная конверсия (а) и один из возможных механизмов генной конверсии (б)

Информация, заключенная в сегменте ДНК, исчезает и заменяется на информацию гомологичного участка другой молекулы ДНК (а). Образовавшийся при рекомбинации гетеродуплексный участок репарируется по образцу одной из цепей гетеродуплекса (б)

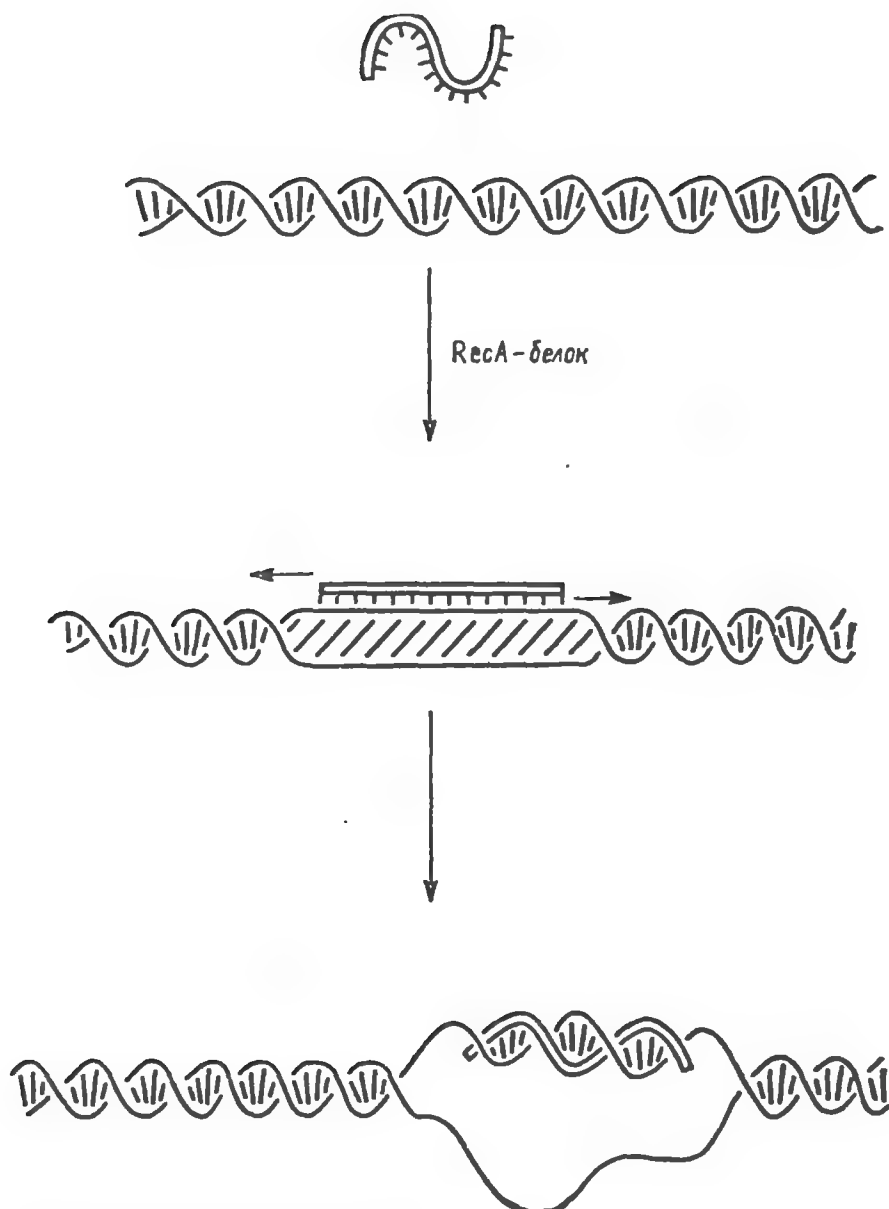


Рис. 58. Образование D-петли

(RecA-белок образует комплекс с одноцепочечной ДНК, который, в свою очередь, взаимодействует с двуцепочечной ДНК. Поиск гомологии осуществляется либо в результате движения одно- и двуцепочечных молекул друг относительно друга, либо в результате повторных актов диссоциации и образования комплекса. Когда гомология найдена, RecA-белок заменяет одну из цепей дуплекса на новую)

одноцепочечной ДНК, необходимый для действия RecA-белка. У *E. coli* ответы на эти вопросы зависят от того, какой работает «путь» рекомбинации, т. е. какие нуклеазы «ассистируют» RecA-белку в ходе рекомбинации. Наиболее эффективным является так называемый RecBC-путь, зависящий, как следует из названия, от продуктов генов *recB* и *recC*, кодирующих две из трех субъединиц одного фермента — экзонуклеазы V или нуклеазы RecBCD. RecBC-путь рекомбинации активируется в бактериях при наличии двуцепочечных концов ДНК, которые могут возникнуть при радиоактив-

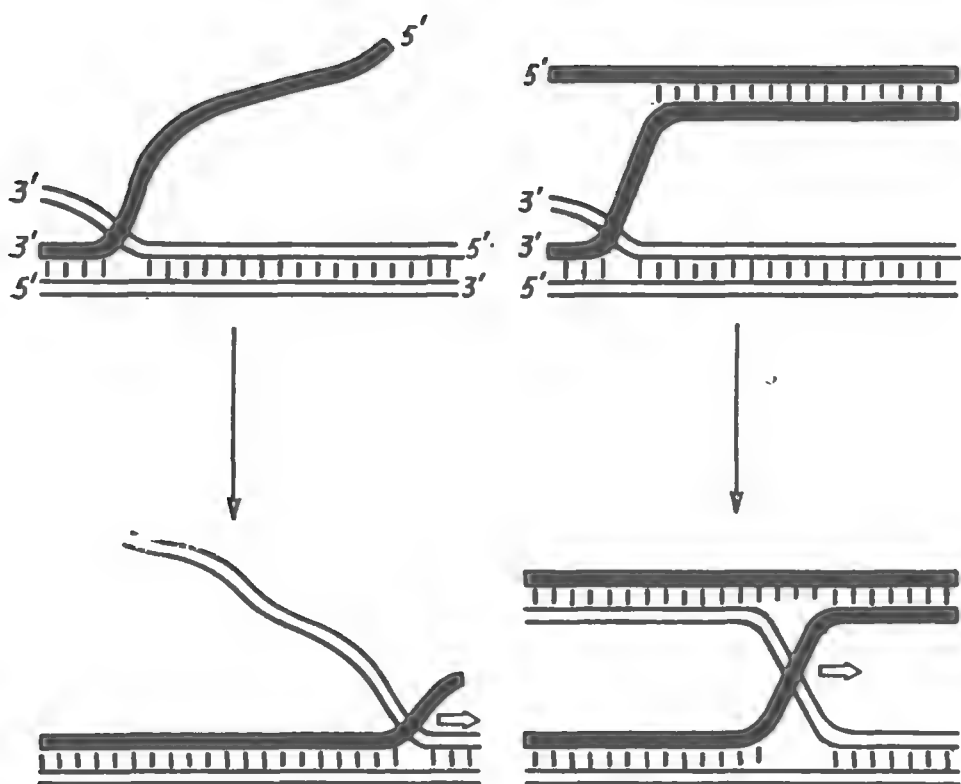


Рис. 59. Направление миграции ветви под действием RecA-белка (Цепи ДНК переносятся начиная с 3'-конца)

ном облучении, при случайном разрыве матричной цепи в репликативной вилке или при переносе ДНК из клетки в клетку при конъюгации (см. гл. V). Нуклеаза RecBCD узнает свободный конец ДНК, присоединяется к нему и осуществляет следующую любопытную реакцию. Используя энергию гидролиза АТФ, фермент перемещается вдоль ДНК и, подобно хеликазам, расплетает дуплекс ДНК. В отличие от других хеликаз RecBCD-белок позволяет ДНК ренатурировать, но расплетание происходит с большей скоростью, чем восстановление дуплекса, поэтому по мере продвижения нуклеазы RecBCD вдоль ДНК постепенно образуются все более протяженные однонитевые петли (рис. 60). Если фермент встречает определенную последовательность ДНК, называемую *хи-сайтом*, то он вносит в нее одноцепочечный разрыв. В результате одна из петель превращается в длинный однонитевой «хвост». Нетрудно видеть, что такая структура является идеальной для дальнейшего взаимодействия с RecA-белком (рис. 60). Любопытно, что последовательность *хи-сайта* асимметрична и разрезается нуклеазой RecBCD лишь в том случае, если она «подъезжает» к *хи-сайту* с нужной стороны.

Для успешного протекания рекомбинации необходим SSB-белок, который покрывает образованную в результате действия RecBCD-белка одноцепочечную ДНК и предохраняет ее от действия нуклеазы: без SSB-белка нуклеаза RecBCD *in vitro*, вместо того чтобы осуществлять описанную реакцию, деградирует обе цепи ДНК

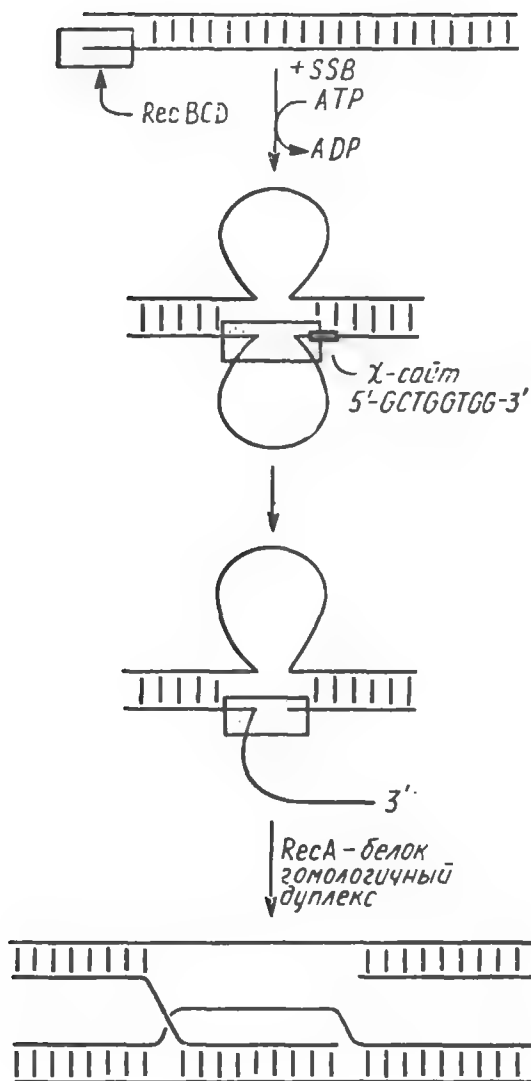


Рис. 60. Схема предполагаемого участия нуклеазы RecBCD (экзонуклеазы V) в инициации гомологичной рекомбинации

(Этот фермент движется по ДНК, расплетая двойную спираль в АТР-зависимой реакции. Одновременно он вновь спаривает комплементарные цепи, но с меньшей скоростью, поэтому по мере продвижения фермента по ДНК возникают все более длинные одноцепочечные петли. Встретив на ДНК определенную последовательность — хи-сайт, нуклеаза вносит в нее одноцепочечный разрыв, создавая тем самым рекомбинационный свободный 3'-конец)

начиная с конца. Кроме того, SSB-белок способствует действию RecA-белка. Мутации по гену *ssb* заметно влияют на рекомбинацию.

Рекомбинация не заканчивается образованием полухиазмы. Эта структура должна быть еще соответствующим образом разрезана. Как это осуществляется у бактерий, до сих пор не вполне ясно. Возможно, за это ответственны те же нуклеазы рекомбинации, которые принимают участие в инициации обмена, например нуклеаза RecBCD. Возможно, существуют специализированные ферменты, специфически разрезающие структуру Холидея. Такие ферменты удалось выделить из клеток, зараженных фагом Т4 или Т7.

В гомологичной рекомбинации могут принимать участие топоизомеразы. Например, создаваемая ДНК-гиразой отрицательная сверхспирализация заметно облегчает образование D-петель, поскольку последний процесс снимает механические напряжения, существующие в сверхспиральных молекулах. Действие топоизомеразы I в принципе позволяет цепям двух рекомбинирующих

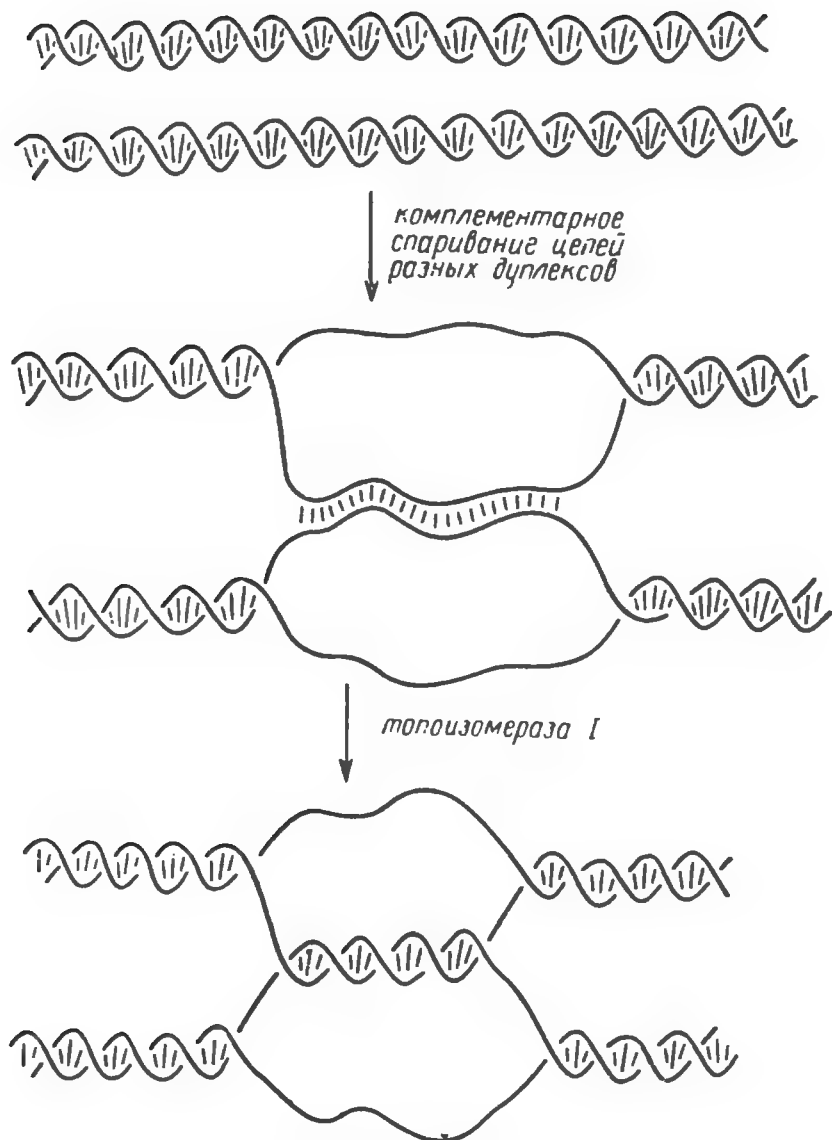


Рис. 61. Возможная роль топоизомеразы I в рекомбинации

(Топоизомераза позволяет ранее не зацепленным цепям, принадлежащим двум гомологичным дуплексам, обвиться друг вокруг друга)

дуплексов комплементарно спариться даже при отсутствии способных обвиться друг вокруг друга свободных концов ДНК (рис. 61).

Рекомбинация и пострепликативная репарация

Одна из ролей гомологичной рекомбинации состоит в репарации повреждений, с которыми не могут справиться описанные в предыдущей главе репаративные системы. Такие повреждения возникают, например, тогда, когда репликативная вилка проезжает через поврежденный участок ДНК до того, как репаративные системы успели устранить повреждение. В этом случае получается, что одна

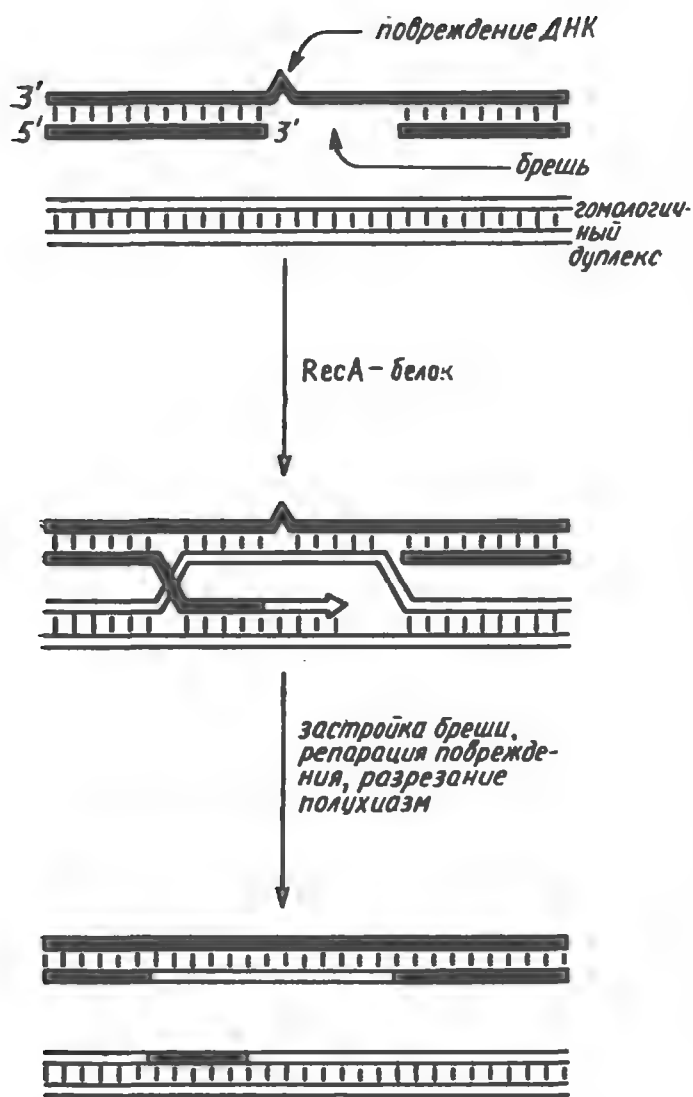


Рис. 62. Пострепликативная репарация

(Если репликация ДНК произошла до того, как из ДНК были удалены дефекты, может возникнуть показанная на рисунке ситуация, когда одна из цепей ДНК повреждена, а во второй имеется брешь, так как поврежденная цепь не может служить полноценной матрицей. Подобные повреждения могут быть исправлены лишь с привлечением гомологичного дуплекса)

разорванная, так и поврежденная цепи ДНК оказываются спаренными с неповрежденными комплементарными цепями, что позволяет их репарацию описанными в предыдущей главе репарационными системами (рис. 62). Таким путем осуществляется пострепликативная, или рекомбинационная, репарация. Аналогичным образом за счет рекомбинации происходит репарация двуцепочечных разрывов ДНК.

Из изложенного следует, что, по крайней мере у бактерий, системы репарации и рекомбинации сложным образом переплетены и связаны между собой. Наличие повреждений ДНК активирует обе

из цепей ДНК дефектна (например, содержит тиминный димер или AP-сайт), а комплементарная цепь не могла быть синтезирована из-за дефекта в матрице и поэтому напротив поврежденного участка остается незастроенная брешь (см. рис. 47). Единственный способ безошибочной репарации такого повреждения — это использовать в качестве эталона второй полученный при репликации дуплекс ДНК, т. е. использовать рекомбинацию для репарации повреждения. У *E. coli* эту задачу способен выполнить RecA-белок вместе с ферментами репарации. Для RecA-белка одноцепочечный участок двуспиральной молекулы ДНК, содержащий повреждение, является «излюбленным» участком связывания. Связавшись с таким местом, RecA-белок вовлекает его в рекомбинационное взаимодействие с гомологичным неповрежденным дуплексом, причем как

системы. С другой стороны, репарация некоторых типов повреждений вообще невозможна, без гомологичного дуплекса ДНК, т. е. может происходить лишь рекомбинационным путем.

Митотическая рекомбинация у эукариот

Рекомбинация может происходить между гомологичными генами соматических клеток многоклеточных или при вегетативном росте одноклеточных эукариот. Частота этой рекомбинации очень невелика, поскольку такая, *митотическая*, рекомбинация может сопровождаться нежелательными последствиями (например, возникновением мозаицизма). Большинство случаев митотической рекомбинации, по-видимому, связаны с репарацией. Действительно, митотическую рекомбинацию можно существенно стимулировать, повредив ДНК, например облучением.

Механизм митотической рекомбинации сравнительно хорошо изучен у дрожжей. На этом объекте показано, что в соответствии с вышеизложенным двуцепочечные разрывы в ДНК на несколько порядков стимулируют рекомбинацию. Стимулирующее действие оказывает не только разрыв, но и двуцепочечная брешь, причем

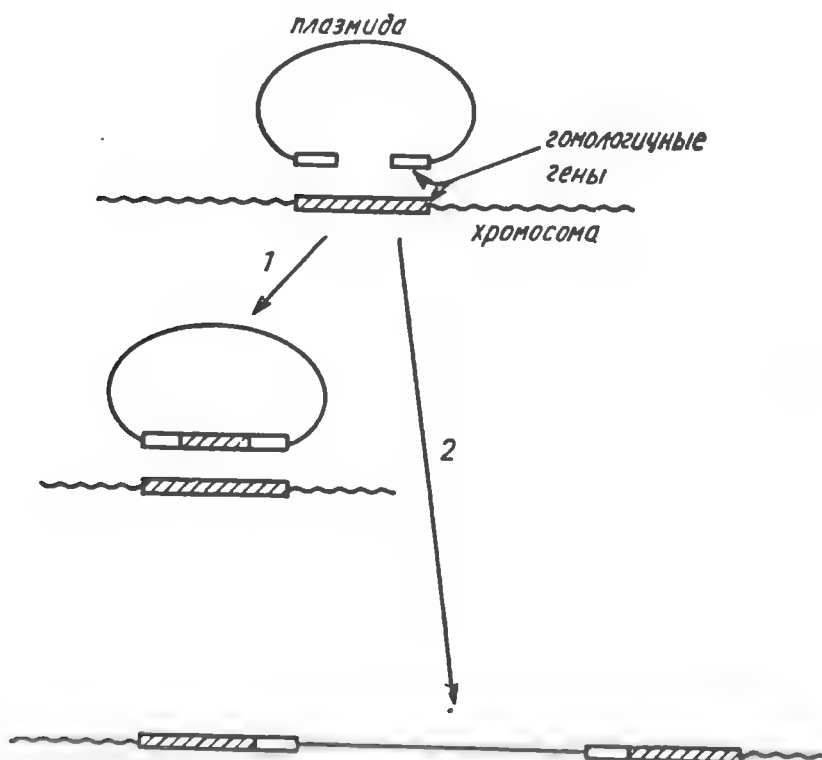


Рис. 63. Один из механизмов геной конверсии

Рекомбинация между гомологичными генами, один из которых разорван двуцепочечной брешью, сопровождается застройкой брешки по образцу ненарушенной копии гена. Конверсия может сопровождаться (2) или не сопровождаться (1) кроссинговером между рекомбинирующими молекулами ДНК

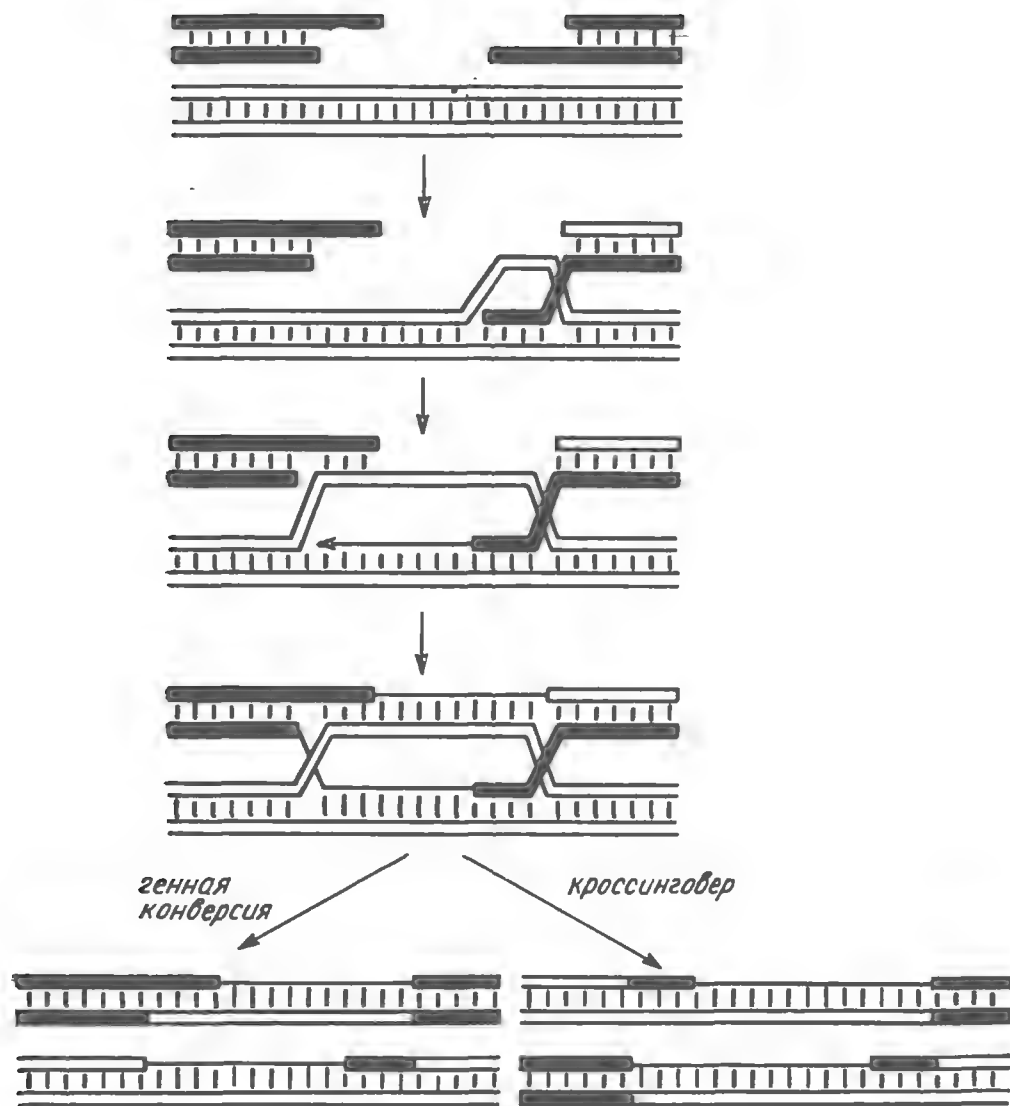


Рис. 64. Модель рекомбинации, инициируемой двуцепочечным разрывом одного из взаимодействующих дуплексов

(Считается, что разрыв может расширяться до бреши за счет действия экзонуклеаз. Согласно этой модели, в ходе рекомбинации возникает промежуточное соединение с двумя холидеевскими структурами. Если обе полухиазмы разрезать одинаково, то рекомбинационное взаимодействие не будет сопровождаться кроссинговером, если же полухиазмы разрезать по-разному, то произойдет кроссинговер фланговых маркеров)

в ходе рекомбинации брешь застраивается по образцу второй рекомбинирующей молекулы. Например, если наблюдать рекомбинацию между введенной в дрожжевые клетки плазмидой, несущей какой-нибудь дрожжевой ген, нарушенный двуцепочечной брешью, и таким же геном в хромосоме (рис. 63), то нарушенный плазмидный ген восстановится по образцу хромосомной копии. Эти данные позволили сформулировать модель рекомбинации-репарации двуцепочечного разрыва. Для этой модели характерно образование двух, а не одной, холидеевских структур. На рис. 64 представлен один из вариантов этой модели. Поскольку двуцепочечная брешь одной из рекомбинирующих молекул ДНК застраивается по образцу

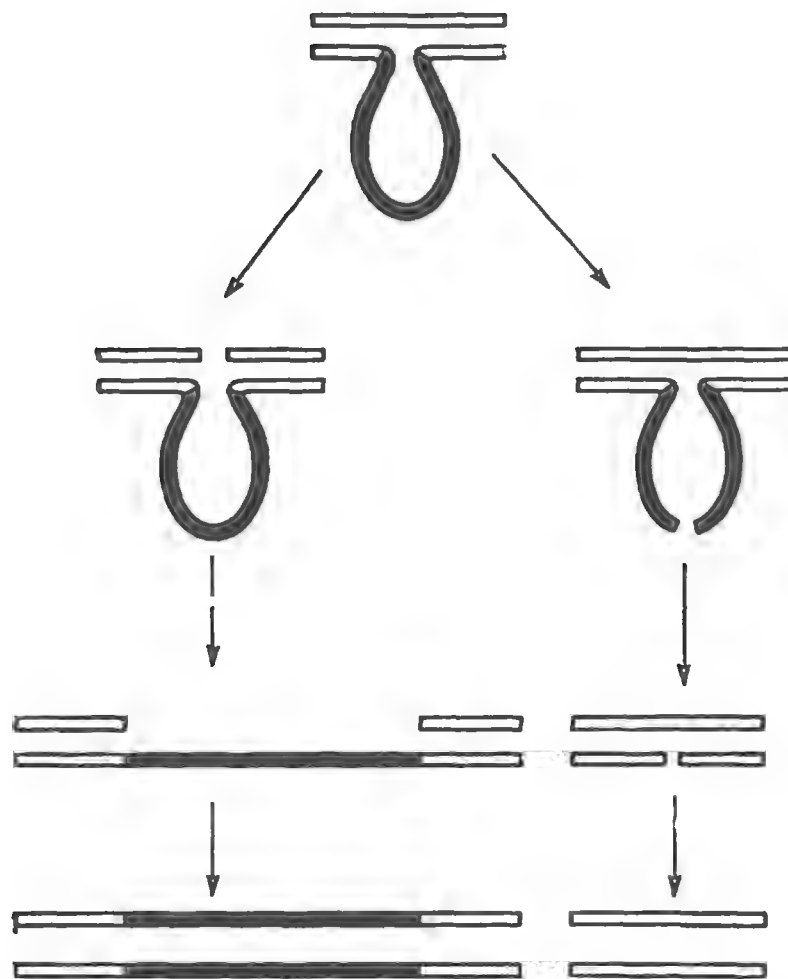


Рис. 65. Генная конверсия протяженных вставок (или делеций)

Если разрыв происходит в молекуле без вставки (слева), то возникает ситуация, неотличимая от репарации двуцепочечной бреши, в результате вставка перемещается в молекулу ДНК, ранее ее не имевшую. Если же разрыв происходит во вставке (справа), то репарация окажется невозможной до тех пор, пока экзонуклеазы не гидролизуют всю вставку и концы разрыва не окажутся гомологичными неразрывному дуплексу. Таким образом, направление конверсии в сторону присутствия вставки или ее отсутствия задается тем, в какую из двух рекомбинирующих молекул ДНК вносятся инициирующий разрыв

неповрежденной молекулы, информация, содержащаяся на участке бреши, полностью теряется и заменяется на информацию неповрежденной молекулы, т. е. репарация двуцепочечной бреши — это еще один механизм генной конверсии. Продукт генной конверсии формально аналогичен результату двойного кроссинговера, но отличается отсутствием второго, реципрокного, продукта рекомбинации. Модель репарации двуцепочечного разрыва объясняет не только случаи конверсии точечных мутаций, но и конверсию протяженных вставок и делеций, которая также часто наблюдается у дрожжей (рис. 65).

Ряд результатов, полученных на дрожжах, подтвержден на других эукариотах. Показано, например, что двуцепочечные разрывы

существенно стимулируют гомологичную рекомбинацию ДНК в культивируемых клетках млекопитающих. Показано также, что репарация таких разрывов сопровождается генной конверсией.

Ферментативный аппарат, осуществляющий рекомбинацию у эукариот, как правило, не известен, хотя ясно, что выявление ферментов эукариотической рекомбинации — дело ближайших нескольких лет. Похоже, что центральную роль в рекомбинации должен играть белок, напоминающий RecA-белок *E. coli*. Действительно, у низшего эукариотического организма *Ustilago maydis* выделен такой RecA-подобный белок — продукт гена *rec1*. Его свойства в целом напоминают свойства RecA-белка, но перенос цепи с одного дуплекса ДНК на другой белок Rec1 начинает с 5'-конца. Любопытное свойство белка Rec1 состоит в том, что в ходе АТР-зависимого поиска гомологии и гомологичного спаривания он, видимо, образует структуру ДНК, в которой цепи не закручены друг вокруг друга (в ограниченной степени это относится и к RecA-белку *E. coli*). Достигается это за счет того, что в комплексе с белком Rec1 ДНК находится частично в обычной В-форме, а частично в левозакрученной Z-форме, т. е. делает виток влево, виток вправо. Поэтому не исключено, что последовательности ДНК, способные легко переходить в Z-форму (альтернирующие пурин-пиримидиновые последовательности), служат участками предпочтительной инициации рекомбинации.

Выделен также белок дрожжей, способный разрезать полухиазмы и по свойствам напоминающий аналогичные белки фагов Т4 и Т7.

Мейотическая рекомбинация

В мейозе у эукариот гомологичные хромосомы, полученные организмом от отца и от матери, расходятся по разным клеткам. Правильное расхождение, при котором гомологи каждой пары расходятся в первом делении мейоза по разным дочерним клеткам, обеспечивается тем, что гомологичные хромосомы узнают друг друга и спариваются на всем своем протяжении (синапс). Две вытянутые друг вдоль друга гомологичные хромосомы образуют в профазе первого деления мейоза особую структуру, так называемый *синаптонемный комплекс* (рис. 66). Молекулярное строение синаптонемного комплекса не ясно, но считается, что эта структура, обеспечивающая гомологичное спаривание хромосом, способствует рекомбинации между ними. Собственно рекомбинация проходит, по-видимому, в специализированных структурах, наблюдающихся на синаптонемном комплексе и называемых *рекомбинационными узелками*. Рекомбинационный обмен, *кроссинговер*, между гомологичными хромосомами в мейозе играет важную роль. Когда перед первым делением мейоза хромосомы конденсируются, то сестринские хроматиды удерживаются вместе, но из-за кроссинговера между несестринскими хроматидами гомологичные хромосомы тоже оказываются

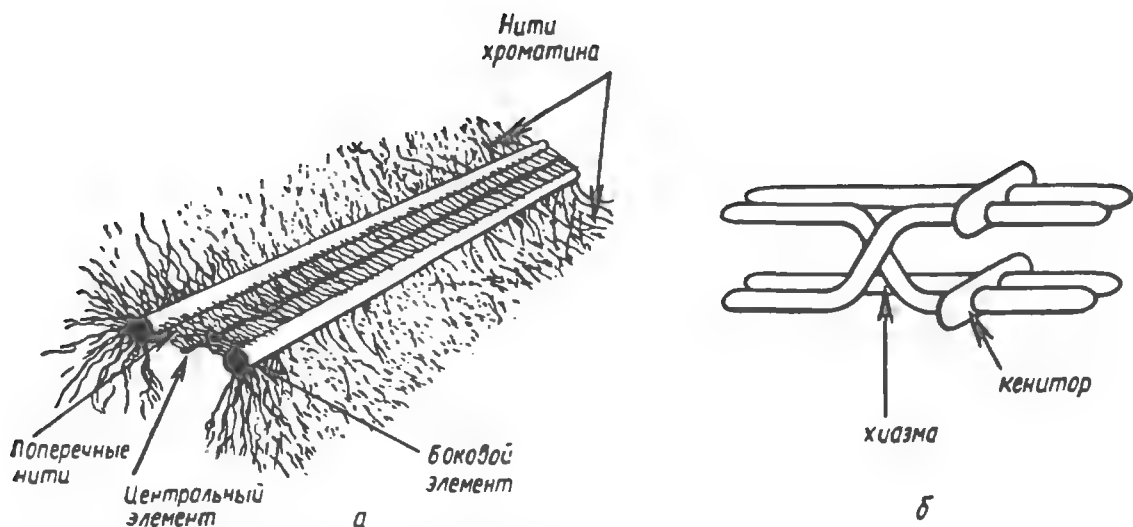


Рис. 66. Мейотическая рекомбинация:

а — схема синаптонемного комплекса, *б* — хиазмы, возникшие вследствие кроссинговера между гомологичными хроматидами, удерживают вместе гомологичные хромосомы, чем обеспечивают их упорядоченную сегрегацию в первом делении мейоза

связанными в участках кроссинговера, называемых *хиазмами* (рис. 66). Такое связывание гомологов обеспечивает их правильную ориентацию на экваторе делящейся клетки и последующее упорядоченное расхождение по дочерним клеткам. Таким образом, возникшие благодаря гомологичной рекомбинации хиазмы выполняют в первом делении мейоза функцию, аналогичную функции центромера в митозе (второе деление мейоза напоминает обычный митоз, упорядоченную сегрегацию хромосом во втором делении обеспечивают центромеры).

Механизм образования синаптонемного комплекса и ферментативный аппарат, обеспечивающий рекомбинацию в мейозе, не изучены, однако большое количество генетических данных, полученных в основном на аскомицетах, указывает на то, что происходит при рекомбинации с ДНК. Мейотическую рекомбинацию сопровождают как генная конверсия, так и постмейотическая сегрегация в том месте, где произошел кроссинговер; значит, этот вид рекомбинации проходит через образование полухиазм и гетеродуплексных участков ДНК. Многие данные по рекомбинации в мейозе можно объяснить с помощью модели рекомбинации-репарации двуцепочечного разрыва, предложенной для митотической рекомбинации, но есть и другая модель, предложенная в свое время специально для объяснения данных по мейотической рекомбинации и лучше объясняющая некоторые детали процесса (рис. 67). Генная конверсия в этой модели является следствием репарации гетеродуплекса в сторону одной из родительских молекул, отсутствие репарации приводит к постмейотической сегрегации. В настоящее время показано, что двуцепочечный разрыв может инициировать рекомбинацию в мейозе, но, как правило, рекомбинация начинается с одноцепочечного разрыва в соответствии с моделью Мезелсона — Рэддинга (рис. 67).

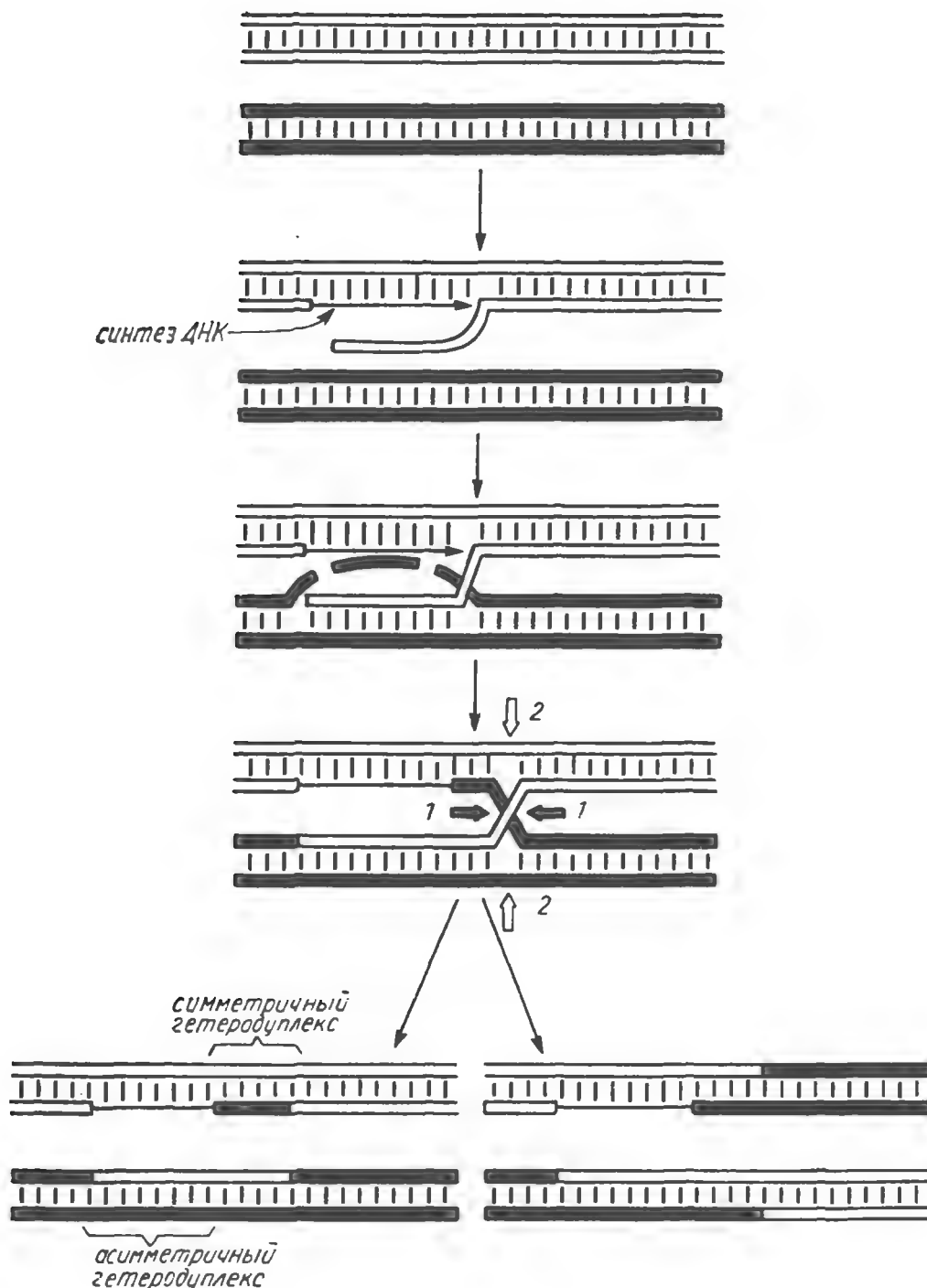


Рис. 67. Модель рекомбинации Мезелсона и Рэддинга

[Согласно этой модели, протяженный гетеродуплексный участок возникает лишь на одной из рекомбинирующих молекул (асимметричный гетеродуплекс), что позволяет объяснить многочисленные генетические данные по рекомбинации в мейозе у аскомицетов. При рекомбинации возникает одна структура Холидея. Если ее разрезать по стрелкам 1, то обмена флангами рекомбинирующих молекул не происходит и наблюдается лишь генная конверсия (если гетеродуплекс репарируется) или постмейотическая сегрегация (слева внизу). При разрезании полу-хназмы по стрелкам 2 происходит кроссинговер (справа внизу)]

Кроме описанной непосредственной «механистичной» функции в мейозе мейотическая рекомбинация играет очень важную эволюционную роль, которая будет рассмотрена позже.

2. СПЕЦИАЛИЗИРОВАННЫЕ СИСТЕМЫ ГОМОЛОГИЧНОЙ РЕКОМБИНАЦИИ

Кроме систем рекомбинации, выполняющих общие функции, например рекомбинацию в мейозе у эукариот или рекомбинационную репарацию у всех организмов, существуют специализированные системы гомологичной рекомбинации, выполняющие ту или иную частную задачу.

Примеры специализированных систем общей рекомбинации можно найти у паразитов, которые с их помощью борются с иммунной системой организмов-хозяев. Так, у гонококков, вызывающих гонорею, имеется много копий генов, кодирующих иммунологически различные варианты пилина — белка, образующего пили (ворсинки) на поверхности этих бактерий и являющегося одним из основных факторов вирулентности (пили обеспечивают прикрепление гонококков к эпителию мочеполовой системы). Экспрессируются лишь те гены, которые находятся в локусах, называемых *pilE*, они и определяют, какой пилин представлен на поверхности клеток. С довольно высокой частотой ген *pilE* утрачивается и бактерия лишается пилей, но за счет рекомбинационного взаимодействия (скорее всего генной конверсии) «остатков» гена *pilE* с каким-нибудь из молчащих генов, расположенных в других местах, в локусах *pilS*, восстанавливается новый ген *pilE* с последовательностью, характерной для гена *pilS*, донора информации при рекомбинации (рис. 68). Таким способом гонококки сравнительно часто меняют иммунологические свойства пилей и не позволяют развиваться эффективному иммунному ответу.

Совершенно аналогичную стратегию использует возбудитель возвратного тифа *Borellia*, который за счет рекомбинации время от времени меняет структуру поверхностного белка, являющегося главным антигеном, распознаваемым иммунной системой человеческого организма.

Рексминационные переключения поверхностных антигенов патогенных микроорганизмов характерны не только для бактерий. Этот способ ускользания патогена от действия защитных средств организма был, собственно, и открыт на эукариотическом болезнетворном агенте — возбудителе сонной болезни *Trypanosoma brucei*. Как и вышеописанные бактерии, трипаносомы имеют множество копий генов так называемого переменного поверхностного гликопротеина, но экспрессируется только одна копия. Обычно экспрессирующаяся копия находится на конце той или иной хромосомы (в теломерной области). Замена экспрессирующейся копии на новую происходит в результате гомологичной рекомбинации, приводящей к обмену теломерами между хромосомами или в результате генной конверсии с участием экспрессирующейся копии и молчащей копии, расположенной в каком-нибудь другом месте генома трипа-

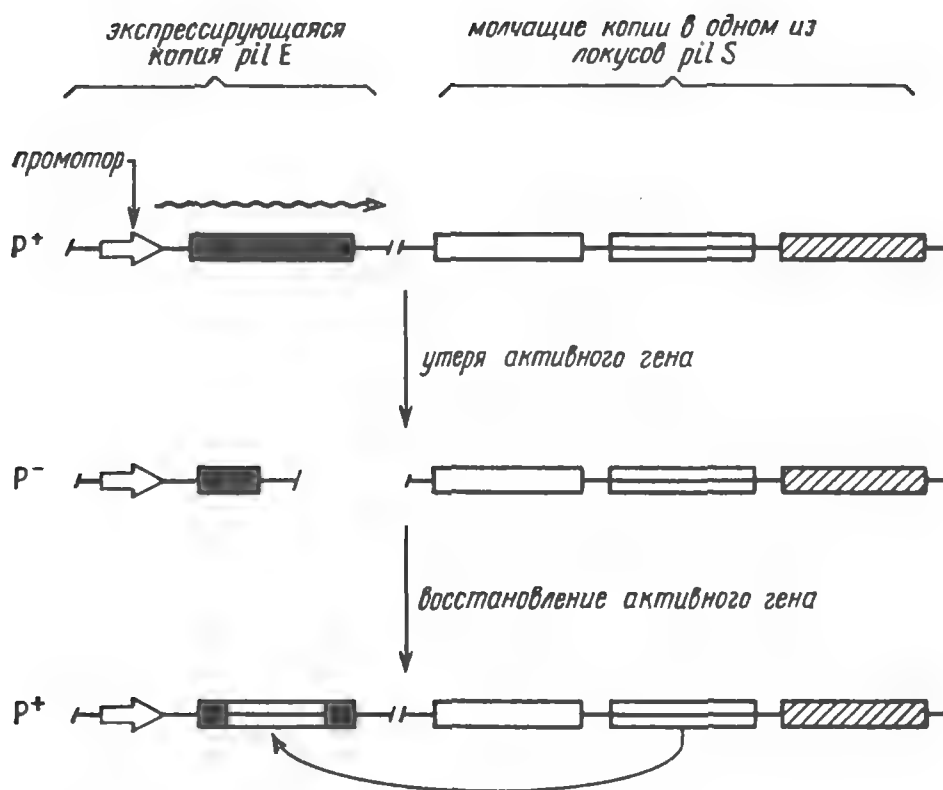


Рис. 68. Механизм антигенных вариаций у гонококков

[Гонококки, экспрессирующие пилин (P^+), с заметной частотой прекращают синтез этого белка (P^-) за счет утраты активного гена в локусе *pil E*. Утрата гена (или его части), видимо, происходит за счет рекомбинации с участием высококонсервативных повторяющихся последовательностей в составе гена. Восстановление активного гена пилина происходит путем рекомбинационного взаимодействия «остатков» исходного гена с гомологичными последовательностями одной из молчащих копий, кодирующих другие варианты пилина и расположенных в локусе *pil S*]

носомы. На этой сравнительно хорошо изученной системе показано, что кроссинговер или конверсия могут не только заменить один ген на другой в предназначенном для экспрессии участке генома, но способны увеличивать разнообразие поверхностных гликопротеинов за счет того, что при рекомбинации возможен перенос не целого гена, а его части и образование гибридного, т. е. нового, варианта гена.

Во всех трех описанных случаях паразитические микроорганизмы используют обычный клеточный рекомбинационный аппарат, а единственная «тактическая хитрость» состоит в организации соответствующих генов: имеется экспрессирующая копия и множество различающихся между собой молчащих. Большое количество последних обеспечивает сравнительно случайный выбор копии для замены и заметную частоту процесса, хотя рекомбинация идет с относительно низкой частотой, которая оказывается достаточной лишь в силу огромных размеров популяций патогенных микроорганизмов. В тех случаях, когда необходима высокая частота рекомбинации, рекомбинационный аппарат клетки должен быть снаб-

жен дополнительными белками. Пример такого рода хорошо изучен на системе смены типа спаривания у дрожжей.

Гаплоидные дрожжи — сахаромикеты — бывают двух типов спаривания: a и α . Клетки разных типов спаривания умеют находить друг друга и образовывать a/α -диплоиды, способные претерпевать мейоз с образованием четырех гаплоидных спор. У некоторых (гомоталлических) штаммов дрожжей происходит эффективное переключение типа спаривания: дрожжи α -типа способны превращаться в дрожжи a -типа, и наоборот. Такое переключение связано со специализированной системой гомологичной рекомбинации. У дрожжей имеется три копии генов, ответственных за тип спаривания. Все три копии расположены на одной хромосоме, но экспрессируются гены лишь одной из них, расположенной в локусе MAT. Слева и справа от MAT-локуса на значительном расстоянии находятся молчащие копии: $HML\alpha$, содержащая информацию, необходимую для того, чтобы дрожжи были α -типа спаривания, и $HMRa$ с информацией, определяющей a -тип спаривания. Последовательность ДНК в MAT-локусе идентична либо левой (MAT α -дрожжи), либо правой (MAT a -дрожжи) молчащей копии. В дрожжах, способных к переключению, с очень высокой частотой происходит замена информации, содержащейся в MAT-локусе, на информацию одной из молчащих копий генов типа спаривания (рис. 69). Показано, что замена происходит путем генной конверсии по механизму репарации двуцепочечного разрыва ДНК. Процесс инициирует специальная сайт-специфическая эндонуклеаза, продукт гена HO, которая вносит разрыв в определенную последовательность ДНК MAT-локуса. Репарация

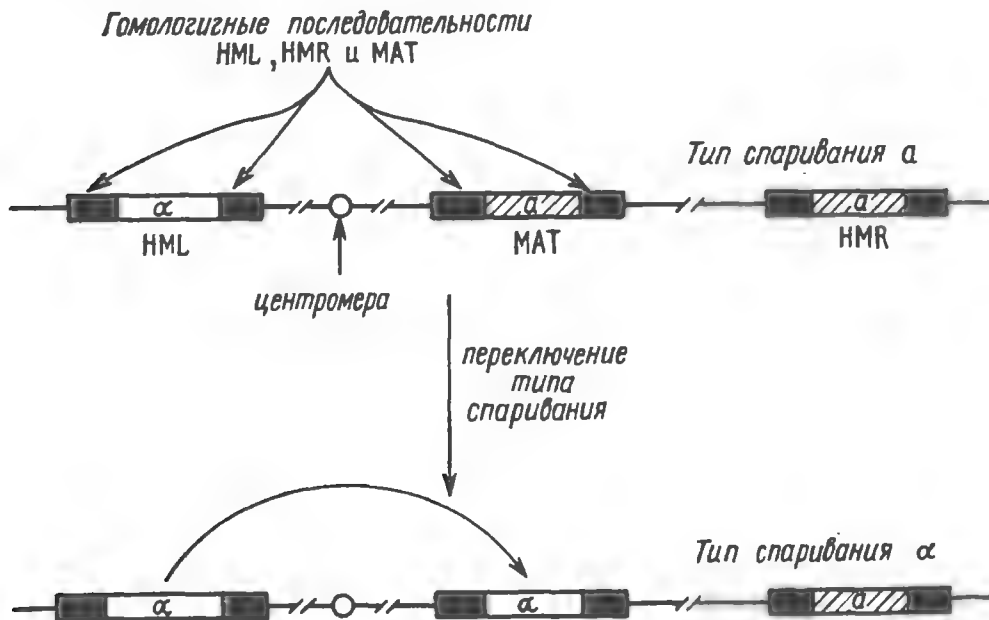


Рис. 69. Расположение локусов $HML\alpha$, MAT и $HMRa$ на третьей хромосоме дрожжей

(Гены типа спаривания не выражаются в локусах HML и HMR, поэтому тип спаривания дрожжевой клетки определяется генами, расположенными в MAT-локусе)

разрыва в гаплоидных клетках может происходить лишь рекомбинационным путем, с использованием в качестве эталона HML-или HMR-копий генов типа спаривания. В одном случае переключения не произойдет, поскольку последовательности ДНК в MAT-локусе и в одной из молчащих копий идентичны, а в другом случае произойдет следующее: экзонуклеазы расширят разрыв до тех пор, пока концы образовавшейся бреши не окажутся гомологичными соответствующим участкам молчащей копии генов, определяющих противоположный тип спаривания. Затем ДНК этих генов используется в качестве матрицы для застройки бреши (см. рис. 64 и 65), т. е. содержащаяся в MAT-локусе информация утрачивается и заменяется на информацию генов другого типа спаривания. В диплоидных a/α дрожжах экспрессируются два MAT-локуса: a и α . В такой ситуации ген HO не работает и никаких переключений не происходит. Таким образом, приспособление дрожжевой клетки для переключения типа спаривания фактически состоит из одного фермента — сайт-специфической эндонуклеазы.

3. САЙТ-СПЕЦИФИЧЕСКАЯ РЕКОМБИНАЦИЯ

Сайт-специфическая рекомбинация — это рекомбинация, которая в отличие от гомологичной не требует протяженных участков гомологии, но для протекания которой необходимы строго определенные последовательности ДНК и специальный ферментативный аппарат. В каждом конкретном случае сайт-специфическая рекомбинация выполняет свою частную функцию, последовательности ДНК для каждого случая также различны, отличаются и ферменты, поэтому имеет смысл рассматривать по отдельности различные системы сайт-специфической рекомбинации. Однако из дальнейшего будет ясно, что в общих чертах механизм сайт-специфической рекомбинации всегда одинаков.

Первая изученная система сайт-специфической рекомбинации — это интеграция фага лямбда в хромосому бактерии-хозяина. Поскольку она описана в главе о вирусах (см. гл. XIII), мы не будем здесь на ней останавливаться, отметим только, что в отличие от рассмотренных случаев хромосомы бактерии и фага не гомологичны, а для рекомбинации необходимы специальные последовательности и специализированный фермент.

Пример другой системы сайт-специфической рекомбинации предоставляет еще один умеренный фаг *E. coli* P1. В отличие от фага λ , P1 в лизогенном состоянии не интегрирует в хромосому клетки, а существует в виде автономной низкокопийной плазмиды. Стабильность наследования таких плазмид зависит от их упорядоченной сегрегации по дочерним клеткам при делении. Механизм сегрегации может нарушаться из-за гомологичной рекомбинации между дочерними молекулами фаговой ДНК после репликации: рекомбинация

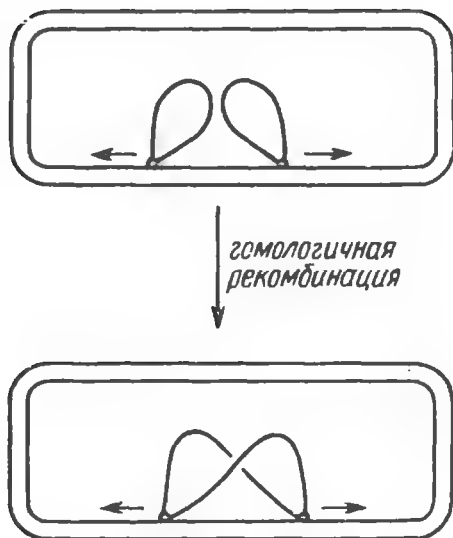


Рис. 70. Гомологичная рекомбинация между двумя кольцевыми репликонами

(Этот процесс может нарушать нормальную сегрегацию репликонов по дочерним клеткам при делении, поскольку образуется димерная молекула, не способная попасть в обе клетки. Для предотвращения этой сложности должен существовать механизм мономеризации репликонов)

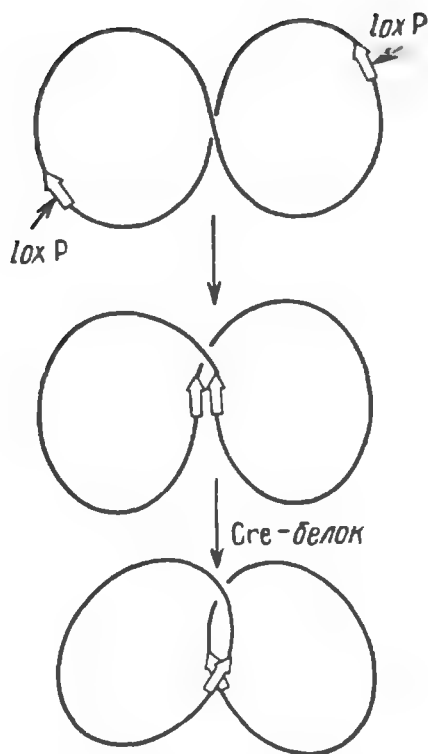


Рис. 71. Механизм мономеризации ДНК фага P1

(ДНК фага P1 содержит один *lox P*-сайт. В результате гомологичной рекомбинации возникает димер с двумя *lox P*-сайтами. Сайт-специфическая рекомбинация между ними, осуществляемая Cre-белком, эффективно мономеризующим ДНК фага P1. В результате реакции могут возникать катенаны, но их «расцепляет» ДНК-гираза)

приводит к образованию димерного кольца, которое при делении попадет лишь в одну из дочерних клеток (рис. 70). Фаг P1 эффективно обходит эту сложность за счет кодируемой им системы сайт-специфической рекомбинации. Продукт фогового гена *cre* с высокой частотой проводит рекомбинацию между двумя специальными последовательностями *loxP*, находящимися на одной молекуле ДНК. ДНК фага P1 содержит одну последовательность *loxP*, но димер, возникший в результате гомологичной рекомбинации между двумя молекулами фоговой ДНК, содержит две последовательности *loxP* в прямой ориентации друг относительно друга. Тут и вступает в игру Cre-белок, вновь разрезающий димер до мономеров (рис. 71). *In vitro* очищенный Cre-белок может осуществлять всю реакцию без каких-либо дополнительных факторов и без доноров энергии. Как и другие белки сайт-специфической рекомбинации, он напоминает топоизомеразы в том отношении, что в ходе рекомбинации вносит разрывы в ДНК и ковалентно связывается с одним из образовавшихся концов, запасая тем самым энергию разорванной связи.

Профаг Р1 не единственный обладатель системы сайт-специфической рекомбинации, необходимой для стабильного наследования. Такие же системы есть у многих других плазмид. В связи с этим необходимо отметить, что к кольцевым репликонам относится также бактериальная хромосома, при упорядоченной сегрегации которой также могут возникать сложности, аналогичные описанным выше для фага Р1 (см. рис. 70). На этом основании можно предположить возможность наличия сайт-специфической системы рекомбинации, мономеризующей бактериальную хромосому перед делением клеток.

Профаг Р1 является обладателем еще одной системы сайт-специфической рекомбинации: в составе фаговой ДНК имеется участок (С-сегмент), который в одних молекулах ДНК имеет одну ориентацию, а в других — другую. За инверсию отвечает продукт фагового гена *cin*, работающий аналогично продукту гена *cre*, с той лишь разницей, что узнает предназначенные для него сайты, фланкирующие С-сегмент, в инвертированной взаимной ориентации. Биологический смысл инверсии сегмента ДНК рассмотрим на примере другого умеренного фага М₁₃, у которого тоже есть инвертируемый сегмент.

Инвертируемый сегмент фага М₁₃ называется G-сегментом, а ген, ответственный за инверсию, — *gin*. G-сегмент имеет размер около 3000 п. о. и фланкирован инвертированными повторяющимися

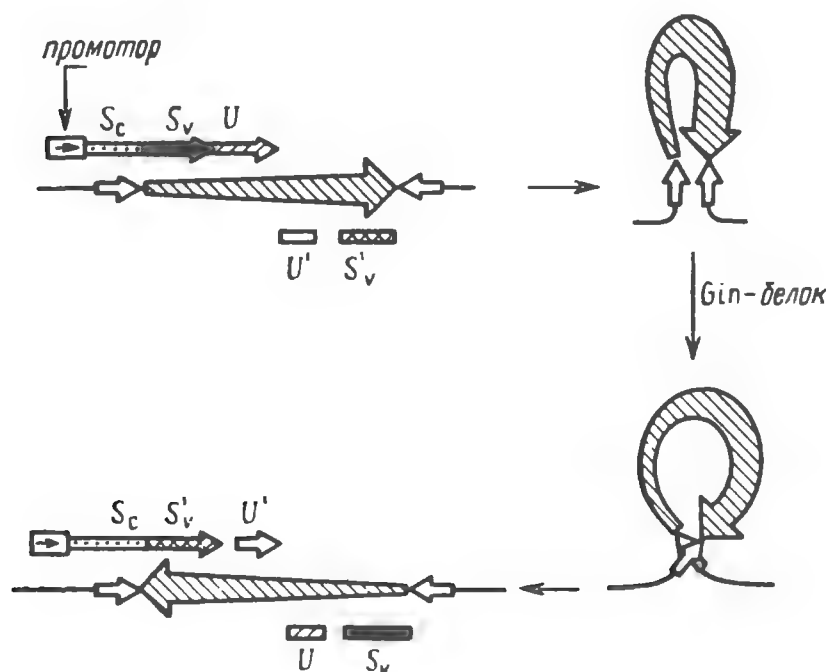


Рис. 72. Инверсия G-сегмента фага М₁₃

(Инверсия происходит благодаря сайт-специфической рекомбинации между двумя инвертированными повторами, окружающими G-сегмент. При одной ориентации сегмента функционируют гены U и S, а при противоположной — гены U' и S'. Так как эти гены ответственны за адсорбцию на бактерии-хозяине, то ориентация G-сегмента определяет круг хозяев бактериофага)

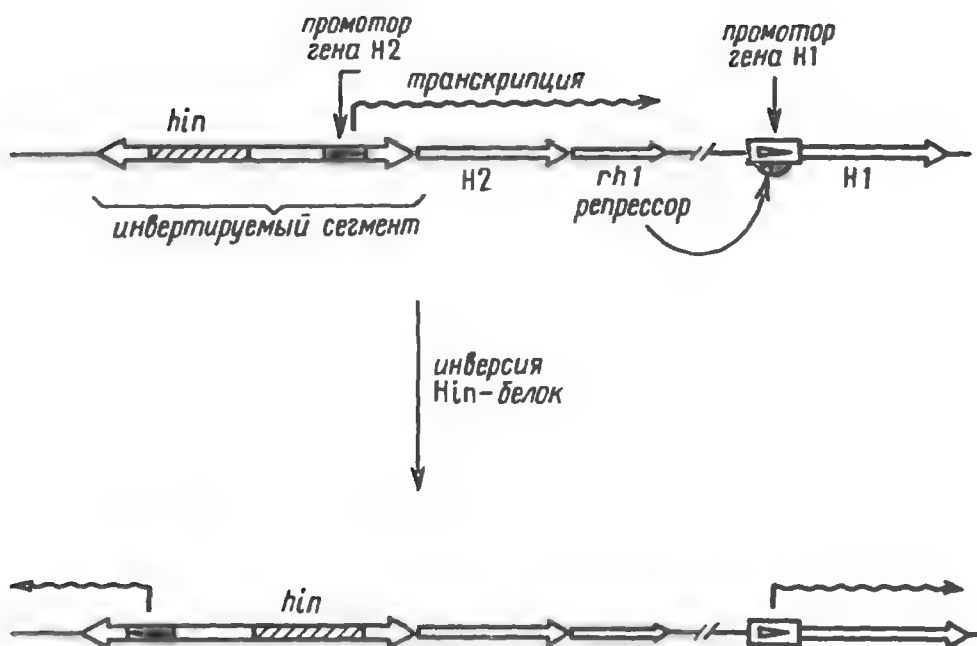


Рис. 73. Смена фаз у сальмонеллы

(Гены H1 и H2 кодируют разные флагеллины. Переключение с одного флагеллина на другой достигается за счет сайт-специфической инверсии сегмента ДНК, содержащего промотор. В одной ориентации промотор направлен в сторону гена H2, который и экспрессируется. Одновременно с флагеллином H2 синтезируется репрессор гена H1. В противоположной ориентации гены H2 и *rh1* не работают, поэтому активен ген H1)

последовательностями длиной около 20 п. о. Сайт-специфическая рекомбинация, осуществляемая G_{in}-белком между этими повторами, приводит к инверсии G-сегмента (рис. 72). И в этом случае G_{in}-белок действует подобно топоизомеразе, самостоятельно осуществляя и разрыв, и перекрестное лигирование молекул ДНК. В составе G-сегмента имеются гены U и S, кодирующие белки отростка фага, причем в одной ориентации G-сегмента работает один набор этих генов, а в противоположной ориентации — другой (рис. 72), поскольку G-сегмент транскрибируется только в одном направлении. Гены U и S определяют круг хозяев бактериофага: при одной ориентации G-сегмента фаг Mu способен адсорбироваться на *E. coli* и *Salmonella typhimurium* и размножаться на этих бактериях, при противоположной ориентации фаг может заражать другие бактерии кишечной группы: *Citrobacter*, *Shigella*, *Erwinia*, *Serratia*. Таким образом, наличие инвертируемого сегмента позволяет фагу Mu существенно расширить круг хозяев. Индивидуальную бактериальную клетку заражает фаговая частица с вполне определенной ориентацией G-сегмента, и при литическом размножении фага эта ориентация преимущественно сохраняется. Но при длительном существовании фага в лизогенных клетках в виде профага происходят множественные акты инверсии, в результате разные лизогенные клетки содержат профаги с различной ориентацией G-сегмента.

Рекомбинационный механизм переключения активности генов используют не только фаги. Аналогичным образом *S. typhimurium*

меняет тип флагеллина, из которого построены жгутики на ее поверхности. Это явление называется *сменой фаз* (рис. 73). Необходимо отметить, что ферменты сайт-специфической рекомбинации, осуществляющие инверсию сегментов фагов Р1 и М_и и смену фаз у сальмонеллы, заметно гомологичны между собой (и отчасти взаимозаменяемы), что свидетельствует об общности их происхождения.

Сайт-специфические рекомбинационные системы есть и у эукариот. Некоторые из них очень напоминают бактериальные, однако биологический смысл рекомбинации в этих случаях, как правило, не ясен. В единственном случае, когда роль сайт-специфической рекомбинации вполне понятна, не известен механизм рекомбинации. Этот случай — сборка из переменных и константных частей функциональных генов иммуноглобулинов и рецепторов Т-клеток.

4. ЭВОЛЮЦИОННАЯ РОЛЬ РЕКОМБИНАЦИИ

Выше уже упоминались функции рекомбинации в клетке. Сайт-специфическая рекомбинация всегда играет ту или иную вполне конкретную частную роль. Гомологичная рекомбинация тоже имеет вполне конкретные непосредственные функции, например обеспечение пострепликативной репарации или осуществление спаривания гомологов, необходимого для правильного расхождения хромосом в мейозе. Но кроме этих непосредственных функций рекомбинация играет очень важную роль в эволюции живой природы.

В первую очередь следует остановиться на роли мейотической рекомбинации при половом размножении организмов. Этот процесс обеспечивает создание разнообразных комбинаций родственных вариантов генов, т. е. обеспечивает комбинативную изменчивость. В результате возникает возможность отбора благоприятных комбинаций генов, т. е. геномов. С другой стороны, благодаря рекомбинации отдельные гены могут оказаться в различном «генетическом окружении», другими словами, ген подвергается давлению отбора в некоем усредненном генетическом окружении. Не будь рекомбинации, естественный отбор мог бы действовать лишь на всю совокупность генов одной хромосомы, но не на отдельный ген. Таким образом, рекомбинация обеспечивает саму возможность отбора высокоадаптивных генов. Здесь же стоит отметить, что гены должны располагаться на хромосомах не случайным образом: близкорасположенные гены редко разделяются рекомбинацией и должны подвергаться коадаптации, поэтому эволюция должна была «позаботиться» о том, чтобы рядом располагались не случайные гены, а гены, имеющие определенное отношение друг к другу.

Можно думать, что широкое распространение полового размножения связано именно с его способностью обеспечивать рекомбинацию, а значит комбинативную изменчивость и отбор высокоадаптивных генов. Даже у бактерий имеются аналоги полового процесса,

позволяющие родственным бактериям обмениваться генетической информацией (см. гл. V).

Рекомбинационные процессы играют также ведущую роль в эволюции строения геномов в целом. Дело в том, что перестройки генетического материала часто можно объяснить рекомбинацией между гомологичными последовательностями, оказавшимися в негомологичном положении (роль таких последовательностей могут выполнять, например, мобильные генетические элементы; см. гл. V). На рис. 81 (см. с. 126) показан один важный частный случай «ошибочной» рекомбинации — неравный кроссинговер. В результате этого процесса генетический материал одной из гомологичных хромосом делится, но в другой хромосоме возникает дупликация. Считается, что такие дупликации играют важную роль в возникновении родственных, но различных генов, поскольку присутствие в геноме «лишних» копий какого-либо гена позволяет им сравнительно свободно изменяться, что, в принципе, может привести к возникновению новых функций белка — продукта гена. По всей вероятности, это один из путей возникновения мультигенных семейств, характерных для геномов высших эукариот и кодирующих белки со сходными, но различными функциями.

Между гомологичными генами одного мультигенного семейства (см. гл. IX) также возможны рекомбинационные обмены, например генная конверсия или неравный кроссинговер. Такие обмены могут иметь ряд любопытных следствий. Некоторые мультигенные семейства, например гистоновые гены, состоят из высокогомологичных генов. Рекомбинационные обмены между ними должны способствовать «унификации» последовательности всех генов семейства, так что такие семейства должны эволюционировать как единое целое, без значительной дивергенции отдельных членов семейства. Напротив, у тех семейств, члены которых сильно дивергировали, рекомбинация может множить разнообразие существующих вариантов, поскольку при обмене между двумя генами может получиться третий, ранее не существовавший вариант. Такие события обнаружены не только в случае специализированных рекомбинационных систем, например в генах поверхностного гликопротеина трипаносом, но и в переменных мультигенных семействах млекопитающих, например среди V-генов иммуноглобулинов и среди генов главного комплекса гистосовместимости.

Таким образом видно, что рекомбинация играет разнообразные, подчас противоречивые роли в эволюции генетического материала.

ЛИТЕРАТУРА

- Р. Б. Хесин. Непостоянство генома. — М.: Наука, 1984.
Ф. Айала, Дж. Кайгер. Современная генетика — М.: Мир, 1987. Т. 1, 2.
Дж. Уотсон, Дж. Туз, Д. Курц. Рекомбинантные ДНК. — М.: Мир, 1986.

Глава V

ПЛАЗМИДЫ И МОБИЛЬНЫЕ ГЕНЕТИЧЕСКИЕ ЭЛЕМЕНТЫ БАКТЕРИЙ

В последние годы стало очевидным, что изменчивость как эу-, так и прокариотических организмов связана не только с точечными мутациями, хромосомными перестройками или описанными рекомбинационными событиями, но и с подвижными, или мобильными, генетическими элементами — сравнительно автономными сегментами ДНК, способными встраиваться в геном клетки-хозяина и вырезаться из него. К мобильным элементам можно отнести и некоторые вирусы — в этом случае возможно перемещение не только в пределах генетического материала одной клетки, но и между клетками (см. гл. XIII). У бактерий перенос генетической информации между клетками могут осуществлять не только вирусы, но и плазмиды, многие из которых могут встраиваться в различные участки генома клетки-хозяина и поэтому тоже могут быть отнесены к мобильным элементам. Плазмиды и мобильные генетические элементы играют существенную роль в эволюции бактерий.

1. БАКТЕРИАЛЬНЫЕ ПЛАЗМИДЫ

Кроме хромосомы у большинства видов бактерий существуют другие способные к автономной репликации структуры — плазмиды. Это двуцепочечные кольцевые ДНК размером от 5 до 0,1 % размера хромосомы, несущие гены, не обязательные для клетки-хозяина, или гены, необходимые лишь в определенной среде. Например, плазмиды (R-факторы) многих клинических штаммов несут устойчивость к антибиотикам, как правило, сразу к нескольким. Другие плазмиды определяют болезнетворность патогенных бактерий, например патогенных штаммов *E. coli*, возбудителей чумы и столбняка. Третьи — определяют способность почвенных бактерий использовать необычные источники углерода, скажем нафталин.

Для своей репликации плазмиды используют репликативную машину клетки-хозяина, однако репликация плазмид происходит независимо от хромосомы. Каждая плазида является самостоятельным репликоном, сама контролирует собственную репликацию и поддерживается в клетке в определенном, характерном для нее числе копий. Для характеристики плазмидных репликонов их принято разбивать на группы несовместимости. Дело в том, что если сходство репликонов столь велико, что система регуляции репликации (или система сегрегации молекул ДНК при делении клетки) не может различить их между собой, то две плазмиды оказываются несовместимыми в одной клетке: после роста клеток в неселективных

условиях в них останется либо одна плазмида, либо другая. Две плазмиды разных групп несовместимости не мешают друг другу и могут стабильно сосуществовать в одной клетке.

Некоторые бактериальные плазмиды (обычно достаточно крупные) способны передаваться из одной клетки в другую, иногда даже в клетку другого вида бактерий (как правило, не слишком далекого). Такие плазмиды называются *трансмиссивными*, и их свойства определяются группой генов, ответственных за перенос (гены *tra*). Трансмиссивные плазмиды кодируют специальные ворсинки, половые пили, которые появляются на поверхности клеток, содержащих плазмиды, и способны специфически связываться с поверхностью бесплазмидных клеток. Последующее сокращение пили притягивает клетки друг к другу и между ними образуется мостик, через который плазмидная ДНК может передаться в новую клетку. *tra*-Гены разных плазмид часто сходны между собой.

Способность передаваться в новые клетки — полезное для плазмид качество (см. ниже), но лишь большие плазмиды могут кодировать сложную систему поверхностных изменений клетки, обеспечивающую конъюгацию. Поэтому многие мелкие плазмиды не могут передаваться сами по себе, но способны к передаче в присутствии трансмиссивных плазмид, используя их аппарат конъюгации. Такие плазмиды называют *мобилизуемыми*.

Механизм передачи ДНК из клетки в клетку состоит в том, что специальный белок узнает определенную последовательность, имеющуюся у трансмиссивных и мобилизуемых плазмид и называемую *ориджином переноса*, вносит в эту последовательность одноцепочечный разрыв и ковалентно связывается с его 5'-концом. Затем цепь ДНК, с которой связан белок, переносится в клетку-реципиент, а неразорванная комплементарная цепь остается в клетке-доноре. Весь этот процесс осуществляют белки, кодируемые *tra*-генами трансмиссивной плазмиды, в частности один из этих генов кодирует специальную хеликазу, которая в АТР-зависимой реакции разделяет переносимую в реципиент и остающуюся в доноре цепи ДНК. Клеточный аппарат синтеза ДНК достраивает одиночные цепи и в доноре и в реципиенте до дуплексов. Белок, сидящий на 5'-конце перенесенной цепи, видимо, способствует замыканию плазмиды в реципиентной клетке в кольцо (таким образом, этот белок напоминает по свойствам топоизомеразы I-го типа и родственные ферменты, например А-белок фага ФХ174; см. гл. XIII).

У некоторых бактерий, в особенности грамположительных, существует процесс естественной трансформации (см. раздел 4 этой главы). Находясь в особом, компетентном, состоянии, эти бактерии способны получать ДНК, оказавшуюся в среде (например, ДНК из погибших клеток), в частности плазмидную ДНК. Это еще один путь перемещения плазмид из клетки в клетку. При трансформации грамположительных бактерий в клетку проникает лишь одна линейная цепь ДНК. Поэтому для восстановления кольцевого плаз-

мидного репликона часто необходимо, чтобы в трансформированную клетку попала олигомерная форма плазмиды; замыкание в кольцо в этом случае происходит за счет гомологичной рекомбинации между повторенными последовательностями олигомера.

2. IS-ЭЛЕМЕНТЫ И ТРАНСПОЗОНЫ БАКТЕРИЙ

IS-элементы (от англ. insertion sequences — последовательности-вставки) — это сегменты ДНК, способные как целое перемещаться из одного участка локализации в другой (рис. 74). IS-элементы содержат лишь те гены, которые необходимы для их собственного перемещения — *транспозиции*. Кроме того, IS-элементы имеют особую последовательность на концах, как правило, *инвертиро-*

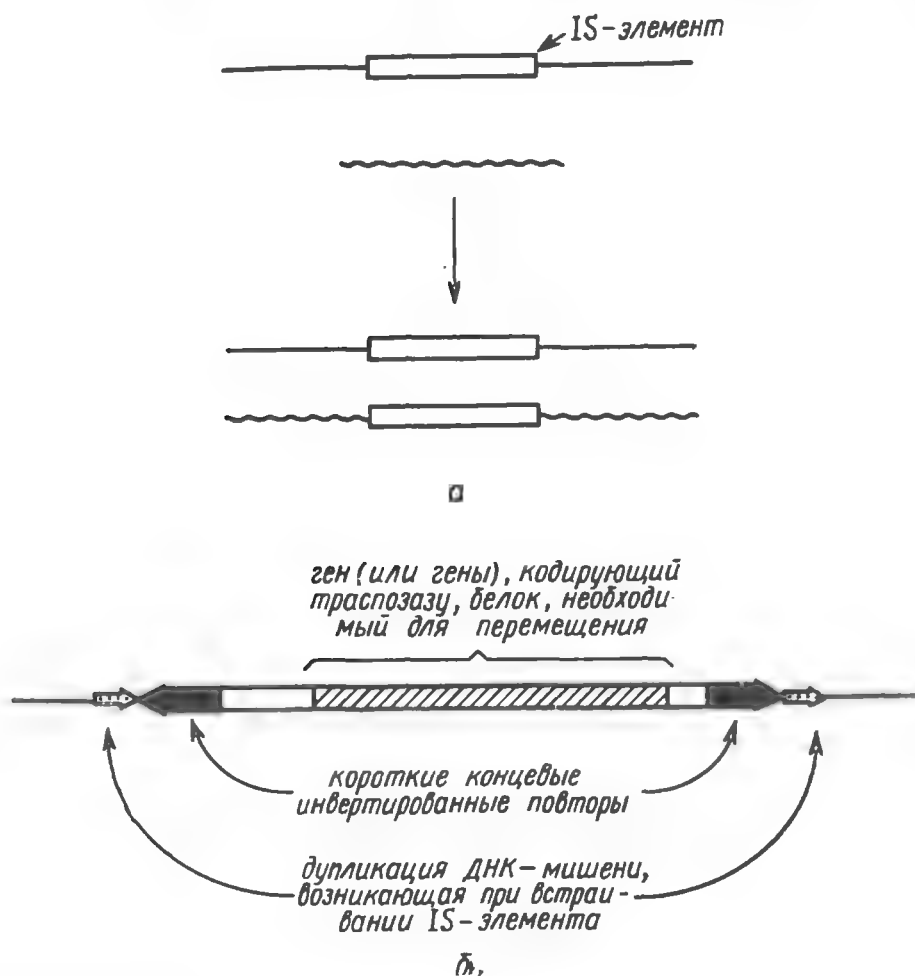


Рис. 74. IS-элементы

[Бактериальные IS-элементы — это сегменты ДНК, способные как целое перемещаться с одного места локализации на другое (а). IS-элементы (б), как правило, кодируют белок (белки), необходимый для их перемещения (транспозиции). По концам элемента расположены инвертированные повторы, необходимые для его транспозиции. У разных IS-элементов длина концевых повторов варьирует от 8 до 40 п. н. Повторы могут быть совершенными или несовершенными. При встраивании в новый участок локализации IS-элементы вызывают небольшую дупликацию ДНК-мишени, часто 5 или 9 п. о. Дублицированные последовательности окружают IS-элемент на новом месте с двух сторон. Разные IS-элементы имеют длину от ~750 до ~1600 п. о.]

ванные повторы. При встраивании в новую последовательность ДНК IS-элементы вызывают небольшую дупликацию: дублированный участок с двух сторон фланкирует встроившийся IS-элемент (рис. 74).

Транспозонами (Тп-элементами) называют сегменты ДНК, обладающие теми же свойствами, что и IS-элементы, но содержащие, кроме того, гены, не имеющие непосредственного отношения к транспозиции. Транспозоны могут нести гены устойчивости к антибиотикам, гены токсинов или гены дополнительных ферментов клеточного метаболизма. В общем, для транспозонов характерны те же гены, которые встречаются в плазмидах. Более того, нередко присутствие в составе плазмиды того или иного гена обусловлено наличием в последовательности плазмидной ДНК соответствующего транспозона (см. раздел 3 этой главы). Транспозон может быть устроен так же, как IS-элемент, но с дополнительным геном. Однако важно отметить, что часто два IS-элемента, оказавшиеся поблизости друг от друга, способны перемещаться вместе, одновременно перенося заключенный между ними сегмент ДНК. Таким образом, два расположенных рядом IS-элемента могут образовать транспозон. Транс-

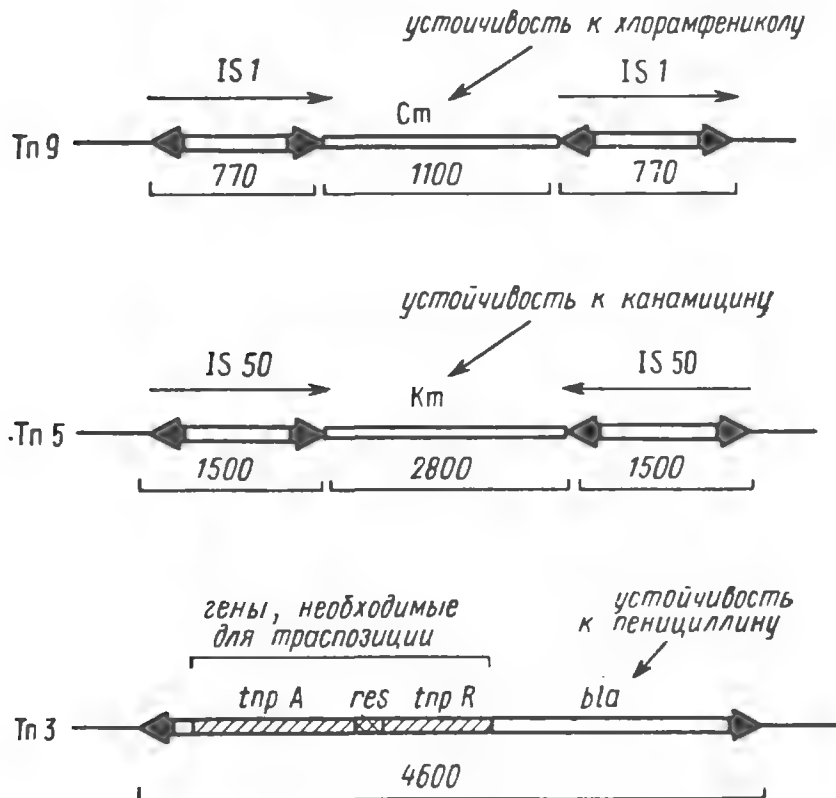


Рис. 75. Некоторые бактериальные транспозоны

[Составные транспозоны состоят из гена, определяющего какой-либо признак (например, устойчивость к антибиотику), окруженного IS-элементами в прямой (Тп9) или инвертированной (Тп5) взаимной ориентации. Ряд транспозонов имеют другое строение и напоминают IS-элемент с дополнительным геном. Примером может служить Тп3]

позоны различной структуры показаны на конкретных примерах на рис. 75. Так как на концах IS-элементов находятся инвертированные повторы, то вне зависимости от взаимной ориентации IS-элементов, образующих тот или иной транспозон, повторы, фланкирующие транспозон в целом, всегда будут инвертированными. Эти повторы необходимы для перемещений транспозона как целого.

Транспозоны и IS-элементы ответственны за целый ряд генетических явлений у бактерий. Встраивание мобильного элемента в какой-либо ген может привести к его инактивации. Кроме того, некоторые IS-элементы и транспозоны вызывают генетическую нестабильность поблизости от места своей локализации: в окрестностях элемента заметно повышается частота делеций и инверсий, причем одна из границ перестройки всегда совпадает с одним из концов IS-элемента, автономного или в составе транспозона. Мобильные элементы способны вызывать также транслокации. Действительно, два расположенных на некотором расстоянии друг от друга IS-элемента можно рассматривать как транспозон и такие «транспозоны» действительно способны к перемещению как одно целое, хотя по крайней мере для некоторых IS-элементов показано, что частота перемещения такой составной структуры быстро падает с увеличением расстояния между фланкирующими IS-элементами.

Очень важное свойство мобильных элементов бактерий состоит в том, что они способны вызывать слияние репликонов, или образование *коинтегратов*. Например, плазмида, в состав которой входит

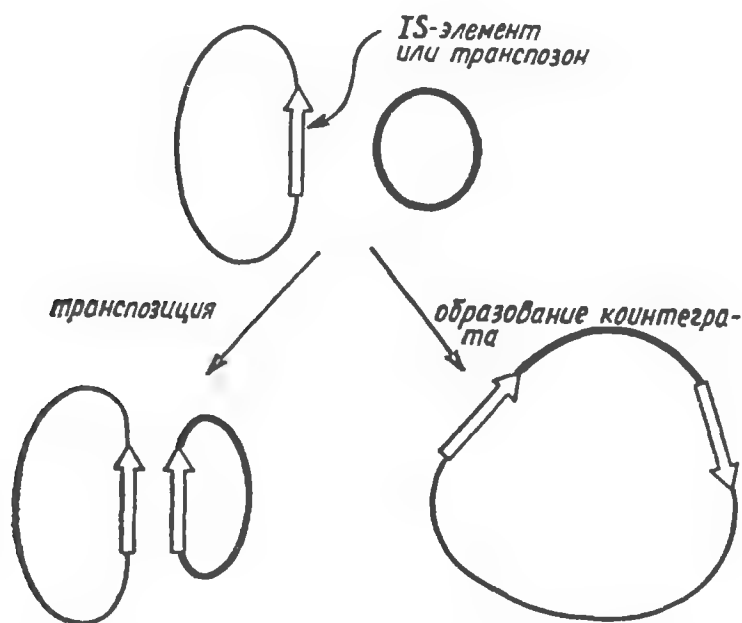


Рис. 76. Две основные реакции, осуществляемые IS- и Tn-элементами

(Мобильные элементы могут не только перемещаться с репликона на репликон, но и вызывать слияние репликонов, образовывать коинтеграт. При этом последовательности двух исходных репликонов оказываются разделенными последовательностью мобильного элемента)

транспозон, может образовать коинтегра́т с другой плазмидой, при этом две плазмиды сливаются в одну, а на границах между последовательностями двух исходных плазмид оказывается по копии транспозона (рис. 76).

Механизм перемещения мобильных элементов бактерий

Как видно из изложенного, при образовании коинтегра́та возникает дополнительная копия мобильного элемента. Аналогичным образом при перемещении с места на место IS-элементы и транспозоны также «размножаются»: копия элемента, как правило, не только появляется на новом месте, но и остается на прежнем (рис. 76). Другими словами, в ходе транспозиции мобильные элементы каким-то образом реплицируются. Как будет показано ниже, для одних транспозонов этот эффект — кажущийся, а в случае других транспозонов действительно происходит репликация мобильного элемента.

Молекулярный механизм транспозиции может быть различным у разных мобильных элементов, поэтому лучше всего рассмотреть его на конкретных примерах. Достаточно изучен в этом отношении бактериофаг *Mu*, являющийся, по сути дела, необычным транспозоном. Этот умеренный бактериофаг встраивается в произвольный участок хромосомы бактерии-хозяина. Если происходит индукция профага и начинается его вегетативное развитие, то он размножается, не вырезаясь из хромосомы, за счет повторных актов репликативной транспозиции. Вырезание фаговой ДНК из бактериальной происходит лишь при упаковке в фаговые частицы, когда репликация уже прошла. При репликации фага *Mu* транспозиция происходит с очень высокой частотой, поэтому именно эта система изучена лучше других.

Генетический анализ показал, что для репликации-транспозиции фага *Mu* необходимы активные продукты его генов А и В и интактные концы фаговой ДНК. Все остальные функции в размножении фага выполняют белки клетки-хозяина. Продукты генов А и В образуют фермент, играющий центральную роль в транспозиции, — *транспозазу*.

Продукт гена А специфически взаимодействует с концами фаговой ДНК и вносит одноцепочечные разрывы точно по концам фаговой последовательности: один — в одну цепь ДНК, а другой — в другую. Так как разрывы разных цепей ДНК расположены по разные стороны мобильного элемента, то одинаковые последовательности узнавания транспозазы на двух концах транспозона должны находиться на разных цепях ДНК. Это объясняет, почему концы мобильного элемента представляют собой инвертированные повторы. Молекулы А-белка связываются с обоими концами транспозона одновременно и удерживают их в едином комплексе. Такой комплекс

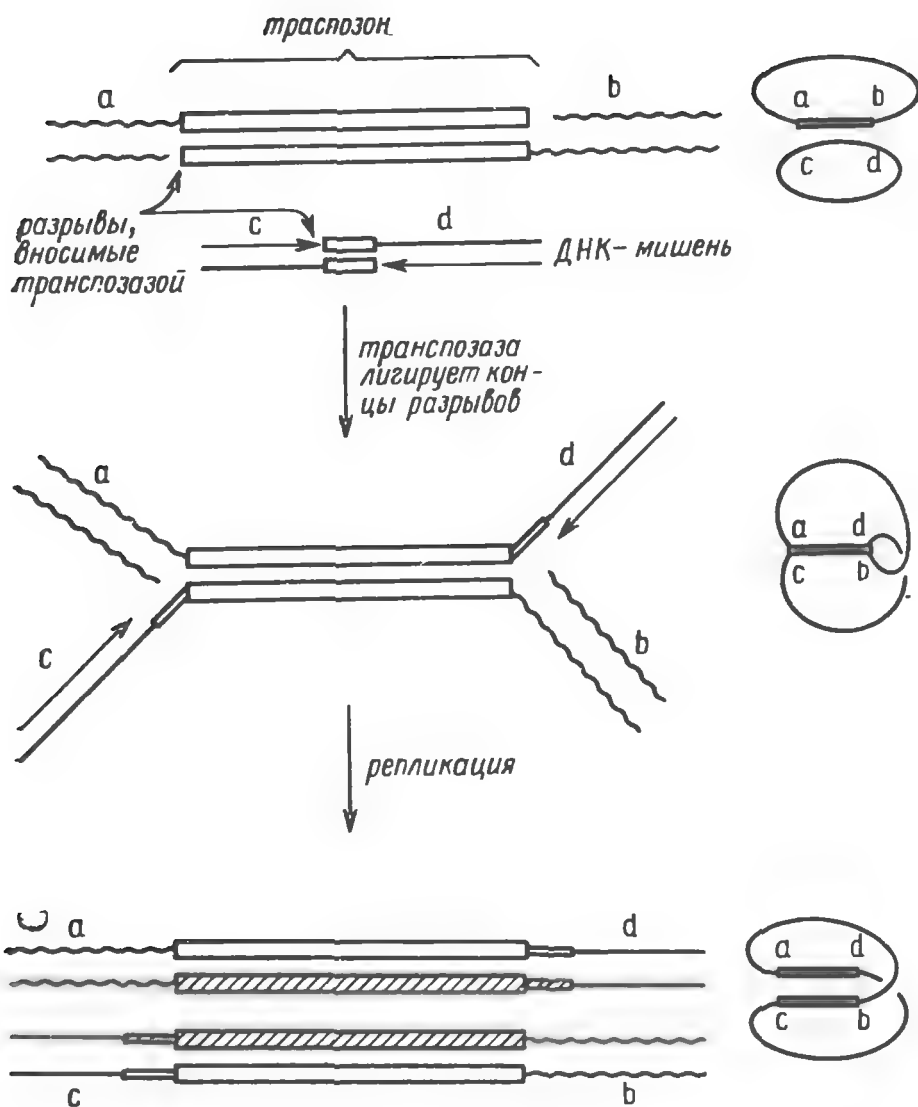


Рис. 77. Механизм репликативной транспозиции

[Транспозаза вносит одноцепочечные разрывы точно по концу транспозона и «косой» двуцепочечный разрыв в ДНК-мишень. Затем транспозаза объединяет концы разрыва мишени с концами транспозона. Возникает промежуточное соединение, напоминающее две направленные навстречу друг другу репликативные вилки. Репликация приводит к удвоению транспозона и, если транспозон и ДНК-мишень находились на разных репликалах, к образованию коинтеграта (слева)]

взаимодействует с ДНК-мишенью (связанной с В-белком). Белки вносят в ДНК-мишень ступенчатый разрыв: комплементарные цепи разрываются на расстоянии 5 п. н. друг от друга. Одновременно транспозаза осуществляет лигирование концов одноцепочечных разрывов с двух сторон от транспозона и концов двуцепочечного разрыва ДНК-мишени, как показано на рис. 77. Фактически этот второй этап реакции состоит в том, что концы разрывов по краям транспозона атакуют фосфодиэфирные связи ДНК-мишени. Эти связи разрываются и одновременно образуются новые связи ДНК-мишени с концами транспозона. Подобная реакция не требует доноров энергии, поскольку общее количество фосфодиэфирных связей не изменяется.

По крайней мере в случае фага М₁ активность транспозазы ограничивается образованием структуры, показанной на рис. 77, дальнейшие события могут происходить без ее участия. Действительно, эта структура не что иное, как две направленные навстречу друг другу репликативные вилки. Репликация за счет клеточного репликативного аппарата приведет к удвоению мобильного элемента и, если транспозон и ДНК-мишень находились на разных кольцевых молекулах ДНК, к образованию коинтеграта (рис. 77). Следствием сдвига в 5 п. н. между двумя разрывами ДНК-мишени является дупликация этого участка после репликации. В случае образования коинтеграта одна копия дублированного участка граничит с одной копией транспозона, а вторая — со второй. В том случае, если произошло перемещение транспозона с репликона на репликон, дупликация фланкирует с двух сторон новую копию транспозона (см. ниже).

Транспозазы некоторых других транспозонов, например Tn3, Tn21 и Tn1000, действуют, по всей вероятности, так же, как транспозаза фага М₁ (отметим, что все эти транспозоны вызывают дупликацию 5 п. н. ДНК-мишени). Так, у транспозона Tn3 имеется два гена, необходимых для его перемещения, *tnpA* и *tnpR* один из которых, *tnpA*, кодирует транспозазу со свойствами, похожими на свойства фермента фага М₁. В результате действия транспозазы Tn3 почти всегда образуются коинтеграты. Чтобы произошло разделение коинтеграта на исходные репликоны (один из которых приобрел новую копию транспозона), необходимо действие продукта гена *tnpR*, называемого *резолвазой* (*resolvase* от англ. *resolution* — разрешение). Оказалось, что этот фермент проводит типичную реакцию сайт-специфической рекомбинации по специальным последовательностям ДНК, имеющимся в составе транспозона и называемым *res*. Реакция разделения коинтеграта показана на рис. 78. Резолваза осуществляет сайт-специфическую рекомбинацию лишь в том случае, если два *res*-сайта находятся на одной молекуле ДНК в прямой ориентации друг относительно друга, т. е. резолваза устроена как раз так, чтобы разрезать коинтеграты на исходные репликоны. Механизм действия резолвазы в общих чертах такой же, как и у других ферментов сайт-специфической рекомбинации, например Gin-белка фага М₁₀ или Cre-белка фага P1 (см. гл. IV, раздел 3).

Ясно, что описанный механизм транспозиции годится не для всех транспозонов. Например, «составные» транспозоны, состоящие из какого-либо гена, окруженного IS-элементами в прямой взаимной ориентации, по структуре неотличимы от коинтеграта (ср. рис. 75 и 76) и поэтому не могли бы существовать в виде стабильной структуры, если бы механизм их транспозиции был таким же, как у транспозона Tn3. Далее, многие транспозоны образуют коинтеграты, отличающиеся исключительной стабильностью. П₁ сколько прямые транспозиции этих транспозонов происходят с большей частотой,

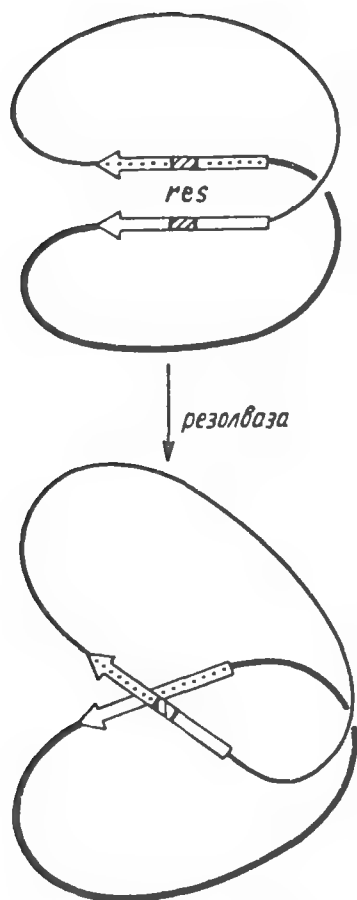


Рис. 78. Роль сайт-специфической рекомбинации в транспозиции

[У транспозонов, перемещающихся с помощью репликативной транспозиции, образование коинтеграта является промежуточной стадией реакции (см. рис. 77). Коинтеграты разрезаются с помощью сайт-специфической рекомбинации между двумя *res*-сайтами, осуществляемой резолвазой]

эти коинтеграты нельзя рассматривать как промежуточную стадию транспозиции.

Перемещение по крайней мере некоторых транспозонов и IS-элементов происходит путем их вырезания с места старой локализации с последующим встраиванием в какое-нибудь другое место того же или нового репликона. Такой механизм называется *консервативным* в отличие от полуконсервативной репликативной транспозиции. В принципе консервативная транспозиция может происходить при незначительной модификации уже описанной схемы, показанной на рис. 77. Действительно, если после образования промежуточной структуры, в которой концы транспозона объединены с концами ДНК-мишени (рис. 77), вместо репликации произойдут разрывы еще двух цепей ДНК, связывающих транспозон со старым местом его локализации, то окажется, что транспозон со старого места вырезался и встроился в новое, а для завершения процесса потребуется лишь ограниченный репаративный синтез ДНК (рис. 79). В этом случае на месте старой локализации транспозона остается двуцепочечный разрыв ДНК. В то же время известно, что перемещение бактериальных транспозонов, как правило, не сопровождается их исчезновением из старого участка локализации. Поэтому предполагают, что вырезание транспозона вызывает «самоубийство донора» — деградацию той конкретной молекулы ДНК, в которой в

результате вырезания транспозона остался разрыв, однако существуют и другие возможности. По-видимому, по описанному или сходному механизму перемещается транспозон Tn5 и ряд IS-элементов, например IS1 и IS903. Все эти (и некоторые другие) мобильные элементы вызывают дупликацию 9 п. н. ДНК-мишени.

По аналогичному механизму перемещается и транспозон Tn10, но в этом случае показано, что вырезавшийся транспозон образует кольцевую молекулу, видимо, способную затем встраиваться на новое место. Образование кольцевой ДНК еще раз указывает на то, что транспозазы одновременно взаимодействуют с двумя концами транспозона или IS-элемента.

Нельзя не отметить, что механизм консервативной транспозиции на первый взгляд уступает полуконсервативному механизму

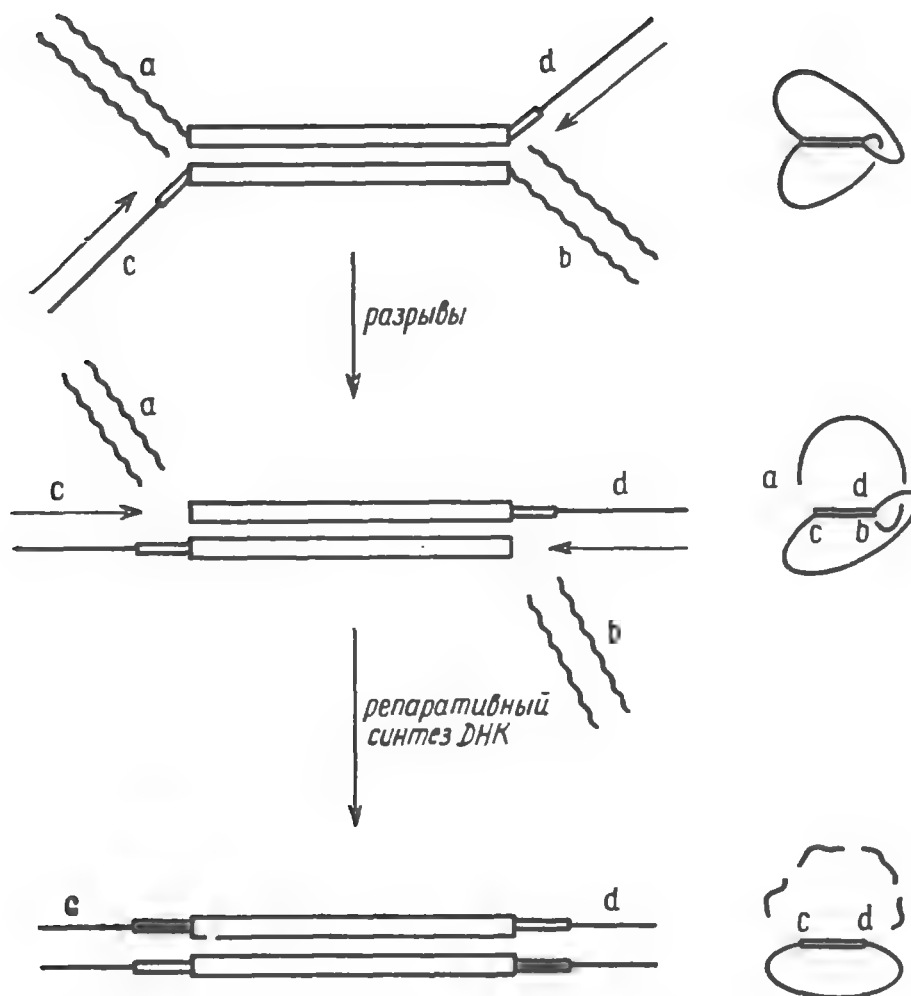


Рис. 79. Консервативный механизм перемещения транспозонов

[Если транспозаза вносит в промежуточное соединение транспозиции (см. рис. 77) два дополнительных разрыва, отделяющих транспозон от его старого окружения, то для завершения транспозиции требуется лишь застройка коротких брешей по обе стороны транспозона. С места старой локализации транспозон вырезается, оставляя после себя двуцепочечный разрыв. Предполагают, что разрыв вызывает «самоубийство донора» — деградацию донорной молекулы ДНК]

репликативной транспозиции. Действительно, в последнем случае происходит не только перемещение мобильного элемента с места на место, но и его размножение, он как бы заражает все новые молекулы ДНК. Однако и транспозиция по консервативному механизму приводит к тому же эффекту. Использующие этот механизм мобильные элементы удивительным образом усовершенствовали его. Оказалось, что синтез транспозазы и активность концов мобильного элемента в реакции транспозиции зависят от состояния метилирования ДНК. Например, у IS10, фланкирующего ген устойчивости к тетрациклину в транспозоне Tn10, в промоторе гена транспозазы и в одном из инвертированных концевых повторов имеются последовательности GATC, которые метилирует метилаза *E. coli dam*. Как уже отмечалось (см. гл. III), сайты метилирования полностью метилированы до репликации ДНК, но после репликации они в течение

некоторого времени полуметилированы. Именно это полуметилированное состояние существенно активизирует и концы IS-элемента, и синтез транспозазы. В результате частота вырезания транспозона из полуметилированной ДНК на несколько порядков выше, чем из ДНК, полностью метилированной. Другими словами, транспозон «чувствует», прореплицировался он или нет. Способность вырезаться сразу после репликации снимает все «недостатки» механизма консервативного вырезания-встраивания. Во-первых, обеспечивается сохранение мобильного элемента на прежнем месте, так как элемент вырезается лишь из одной из двух молекул ДНК, образовавшихся после репликации. Наличие второй копии допускает, в принципе, рекомбинационную репарацию оставшегося после вырезания транспозона двуцепочечного разрыва, хотя не ясно, действительно ли она происходит или имеет место «самоубийство донора». Во-вторых, чувствительность к метилированию обеспечивает дополнительную регуляцию транспозиции: при передаче транспозона в новую бактериальную клетку в составе трансмиссивной или мобилизуемой плазмиды он сразу оказывается в полуметилированном состоянии (при конъюгации передается одна цепь ДНК, которая достраивается до дуплекса в реципиентной клетке), что повышает вероятность его перемещения в другие репликоны клетки-реципиента.

*Механизмы
генетической нестабильности,
вызываемой мобильными
элементами бактерий*

Выше упоминалось, что мобильные элементы вызывают генетическую нестабильность поблизости от участка своей локализации. Эта особенность легко объясняется уже известными нам свойствами IS-элементов и транспозонов бактерий. На рис. 80 показано, что получится при перемещении в пределах одного репликона транспозона типа $Tn3$, т. е. с репликативным механизмом транспозиции. В зависимости от того, как внесены разрывы в ДНК-мишень, получится либо делеция, либо инверсия генетического материала между местом расположения транспозона и мишенью его перемещения. По сути дела, образование делеции напоминает процесс распада коинтеграта, но поскольку одна из образовавшихся молекул ДНК не имеет ориджина репликации, она утрачивается. Если происходит инверсия, то на обеих ее границах оказывается по копии транспозона в инвертированной друг относительно друга ориентации. Таким образом, образование делеций и инверсий характерно для репликативного механизма транспозиций.

Транспозоны, перемещающиеся путем вырезания-встраивания, как правило, подобной нестабильности не вызывают. Способность мобильного элемента стимулировать образование делеций коррели-

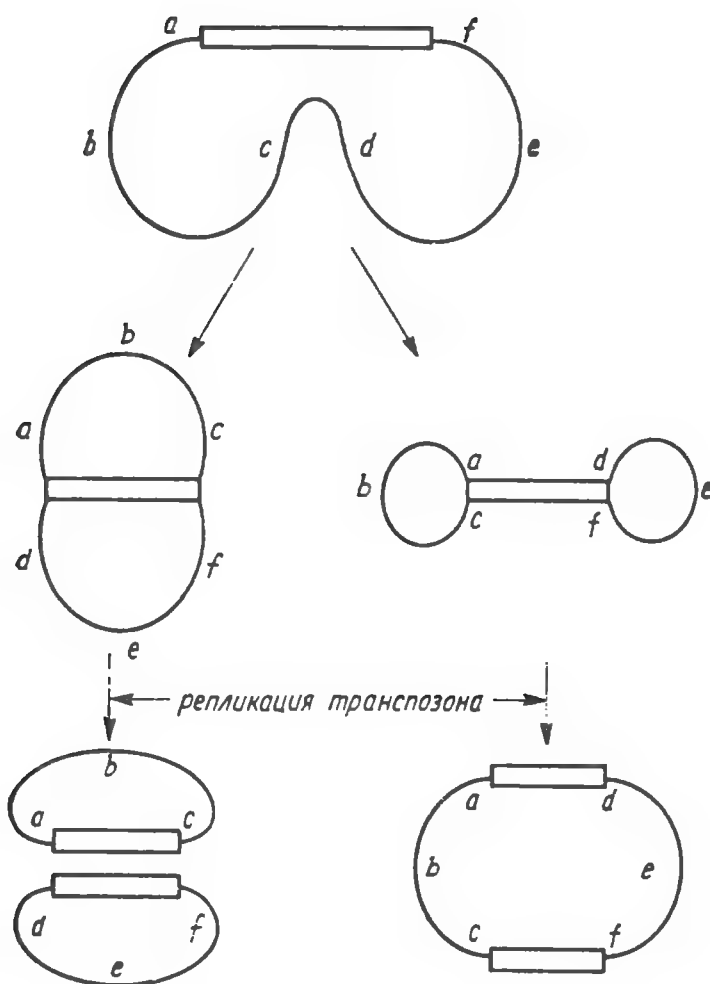


Рис. 80. Механизм генетической нестабильности, вызываемой мобильными элементами [При «попытке» мобильного элемента осуществить репликативную транспозицию в пределах того же репликона, где он находится, возникают делеции (слева) или инверсии (справа)]

рует с его способностью образовывать коинтеграты. Те транспозоны, которые, подобно Tn5 и Tn10, вообще не образуют истинных коинтегратов, вызывают делеции очень редко, по-видимому, за счет abortивного процесса вырезания, при котором транспозаза один разрыв делает по концу транспозона, а второй, по ошибке, — в произвольном месте поблизости.

Индукцированное IS-элементами или транспозонами образование делеций или инверсий не удаляет мобильный элемент, послуживший их причиной, поэтому нестабильность участка локализации мобильного элемента сохраняется.

Бактериальные мобильные элементы с небольшой частотой могут исчезать из места своей старой локализации (этот процесс, судя по всему, не имеет непосредственного отношения к транспозиции). Например, если встраивание транспозона инактивировало какой-либо ген, то его вырезание приводит к реверсии. Такие ре-

версии стабильны, поскольку мобильный элемент полностью утрачивается. На основании полного восстановления активности мутантных генов при вырезании IS-элементов и транспозонов предполагают, что вырезание происходит абсолютно точно и что дублированный участок ДНК-мишени снова оказывается представленным одной копией. В принципе, такой результат может быть достигнут за счет гомологичной рекомбинации между окружающими транспозон дублированными последовательностями или за счет ошибки репликации, вызванной наличием на концах транспозонов инвертированных повторов. Возможно, что последний механизм объясняет и случаи «почти точного» вырезания, когда мобильный элемент исчезает не бесследно, а оставляет после себя измененную последовательность ДНК.

3. ПЛАЗМИДЫ И МОБИЛЬНЫЕ ЭЛЕМЕНТЫ БАКТЕРИЙ И КОНЦЕПЦИЯ «ЭГОИСТИЧНОЙ» ДНК

Какова функциональная роль плазмид и мобильных элементов бактерий? Ниже будет рассмотрена существенная роль этих структур в эволюции бактерий, но эволюционные, т. е. отдаленные, преимущества вряд ли могут объяснить поддержание в бактериальных клетках мобильных элементов и плазмид в тех случаях, когда они не приносят непосредственных селективных выгод. Так, например, если считать, что в клетке поддерживается только функционально необходимый генетический материал, непонятно, почему плазмиды, несущие гены устойчивости к антибиотикам, встречаются не только в клинике, где эти антибиотики применяют, но и в других местах обитания, лишенных подобного селективного давления. Совсем непонятно, почему существуют плазмиды, вообще не приносящие никаких непосредственных преимуществ содержащим их клеткам, и IS-элементы.

Подобные парадоксы можно разрешить, вспомнив, что и плазмиды и мобильные генетические элементы обладают сравнительной автономией от основной массы генетического материала, и поэтому их можно рассматривать как своего рода организмы, обитающие в особой, генетической, среде. Таким образом, можно рассматривать функции плазмид, IS-элементов и транспозонов не с точки зрения преимуществ, которые они несут бактериям-хозяевам, а с точки зрения их самоподдержания в бактериальных популяциях, другими словами, можно считать автономные элементы прокариот «эгоистической» ДНК, обеспечивающей в первую очередь собственное размножение. В этом смысле мобильные элементы и плазмиды непосредственно примыкают к вирусам, «эгоистические» тенденции которых очевидны.

Мобильные элементы и плазмиды используют ряд стратегий,

чтобы избежать элиминации из популяции бактерий, которым они не несут непосредственных преимуществ. Первый очевидный способ — приносить минимальный вред хозяевам. Этот способ используют многие транспозоны, ограничивая частоту своих перемещений очень незначительной величиной (высокая частота транспозиции привела бы к появлению большого количества инсерционных мутаций и к накоплению не нужной клетке ДНК транспозона). Ограничение частоты транспозиции может происходить за счет крайне низкой конститутивной активности гена транспозазы (как, например, в случае Tn10), или за счет действия репрессора, подавляющего синтез транспозазы (в случае Tn3), или блокирующего необходимые для транспозиции инвертированные повторы на концах мобильного элемента (как, по-видимому, происходит в случае Tn5). Любопытно, что у Tn3 и родственных транспозонов репрессором транспозиции служит резолваза: связывание этого белка с *res*-сайтом подавляет транскрипцию гена транспозазы и направленного в противоположную сторону собственного гена резолвазы (см. рис. 75). Естественно, частота транспозиции повышается в «выгодных» для мобильного элемента случаях. Например, если транспозон, активность которого подавляется репрессором, в составе конъюгативной плазмиды передается в клетку-реципиент, в которой такого транспозона нет, то вероятность его перемещения на новый репликон (например, в хромосому реципиента) временно повысится благодаря отсутствию репрессора транспозиции в клетке-реципиенте. Тот же эффект, как уже отмечалось, будет наблюдаться у мобильных элементов, чувствительных к метилированию ДНК: так как в реципиент переносится одна цепь ДНК, то ее достраивание до двуцепочечной молекулы создаст полуметилированную ДНК, повышающую частоту транспозиции.

Ключевым свойством бактериальных мобильных элементов, обеспечивающим их сохранение, является их способность перемещаться с репликона на репликон. Наличие у бактерий трансмиссивных и мобилизуемых плазмид позволяет транспозонам и IS-элементам не только переходить с плазмиды на плазмиду или из хромосомы на плазмиду, но и путешествовать из клетки в клетку в составе плазмид. Таким путем мобильные элементы могут распространяться в бактериальных популяциях, даже если не приносят своим хозяевам никаких преимуществ. В этой связи следует упомянуть о явлении «иммунитет к транспозиции»; многие транспозоны и IS-элементы значительно чаще перемещаются на новые репликоны, чем на новое место в составе того репликона, в котором они находятся. Молекулярный механизм этого свойства еще не выяснен, но очевидно, что оно способствует распространению мобильного элемента по максимальному количеству репликонов.

Шансы на выживание и широкое распространение мобильного элемента еще более усилятся, если он окажется способным приносить клеткам-хозяевам непосредственные выгоды. Для этого он

должен нести в своем составе какой-либо полезный ген. Именно поэтому из IS-элементов могли образоваться транспозоны. Любопытно проследить эволюцию «составных» транспозонов, в которых два IS-элемента окружают «полезный ген». Довольно понятно, что эти структуры возникли в результате внедрения IS-элемента поблизости от этого гена и последующего перемещения того же элемента в другой участок молекулы ДНК по другую сторону от гена. Образовавшаяся структура может перемещаться как целое, но могут по отдельности перемещаться и образовавшие ее IS-элементы. Но если попавший между IS-элементами ген оказался «удачным» (см. ниже), то получившийся транспозон начинает распространяться как целое, чему способствует постепенное накопление мутаций во фланкирующих его IS-элементах. Накопление мутаций приводит к дивергенции IS-элементов, причем по крайней мере один из них теряет способность к самостоятельной транспозиции. Так произошло, например, в случае транспозонов Tn5 и Tn10, лишь один IS-элемент которых способен к синтезу активной транспозазы и к самостоятельному перемещению. Транспозоны типа Tn3 должны были возникнуть в результате более радикального процесса. Скорее всего сначала происходило перемещение соответствующего IS-элемента в пределах того же репликона, поэтому тот или иной ген оказывался окруженным инвертированными копиями мобильного элемента, а затем происходило делетирование значительной части одного из IS-элементов, так что от него оставался лишь один инвертированный концевой повтор.

Какие же гены оказываются «полезными» и входят в состав мобильных элементов? Это не праздный вопрос, поскольку каждая бактериальная клетка хорошо приспособлена к своей среде обитания и не нуждается в генах, аналогичных тем, которые у нее уже есть и обеспечивают ее адаптацию к среде. С другой стороны, приспособление к совершенно новой среде обитания, по-видимому, требует относительно значительной перестройки генетического материала клетки, включающей, в частности, коадаптацию многих разных генов. Поэтому клетка может получить селективное преимущество за счет приобретения какого-либо гена (в составе транспозона) лишь в том случае, если этот ген сам по себе способен оказаться выгодным для бактерии в определенных условиях, т. е. именно такие гены выгодно иметь транспозонам в своем составе. Действительно, на транспозонах «путешествуют» гены устойчивости к различным бактериальным ядам, в том числе к тяжелым металлам и антибиотикам, гены дополнительных метаболических путей, позволяющие использовать, например, какой-нибудь необычный источник углерода, наконец гены некоторых токсинов, делающие бактерии патогенными и позволяющие им тем самым существенно изменить образ жизни.

Сказанное в равной степени относится и к плазмидам, что естественно, поскольку большинство полезных для клетки-хозяина

плазмидных генов находится в составе транспозонов. Однако стратегия выживания плазмид, как и в случае транспозонов, основывается не только на том, чтобы приносить пользу клеткам-хозяевам. Прежде всего сохранению плазмид в популяциях бактерий способствует высокая стабильность их наследования. Она достигается точным контролем репликации и сегрегации плазмид по дочерним клеткам (см. гл. II), но многие плазмиды имеют еще дополнительные приспособления для повышения стабильности наследования. Так, большая конъюгативная плаزمида *E. coli*, называемая F-фактором, имеет гены, блокирующие деление клетки, если в ней находится только одна копия плазмиды, т. е. в ситуации, способной привести к утере плазмиды (этот механизм можно назвать «только со мной»). Другая плазмиды, R1 (кодирующая устойчивость к нескольким антибиотикам), достигает той же цели другим путем. Один из генов этой плазмиды кодирует токсичное для клетки вещество, а другой ген — короткоживущий ингибитор этого вещества, так что наличие плазмиды для клетки не вредно. Но если в результате ошибочной сегрегации одна из клеток после деления теряет плазмиду, то накопившийся в ней токсин приведет к гибели этой клетки, поскольку прекратился синтез ингибитора (механизм «без меня — убью!»). Усовершенствованный вариант последнего механизма используют колициногенные плазмиды (ColE1, ColA и др.), кодирующие бактерицидные белки-колицины и одновременно придающие клеткам свойства иммунитета к их действию. В результате колицины будут убивать лишь бесплазмидные клетки: как клетки, потерявшие плазмиду, так и те, которые ее не имели вовсе. Этот случай интересен тем, что является пограничным: колициногенные плазмиды приносят содержащим их клеткам явную выгоду, уничтожая бесплазмидных конкурентов, но, с другой стороны, селективное преимущество клеток с плазмидой создается явно искусственно; этот класс плазмид не способствует лучшему приспособлению бактерий к среде обитания, но лишь увеличивает шансы содержащих их клеток в конкуренции с себе подобными.

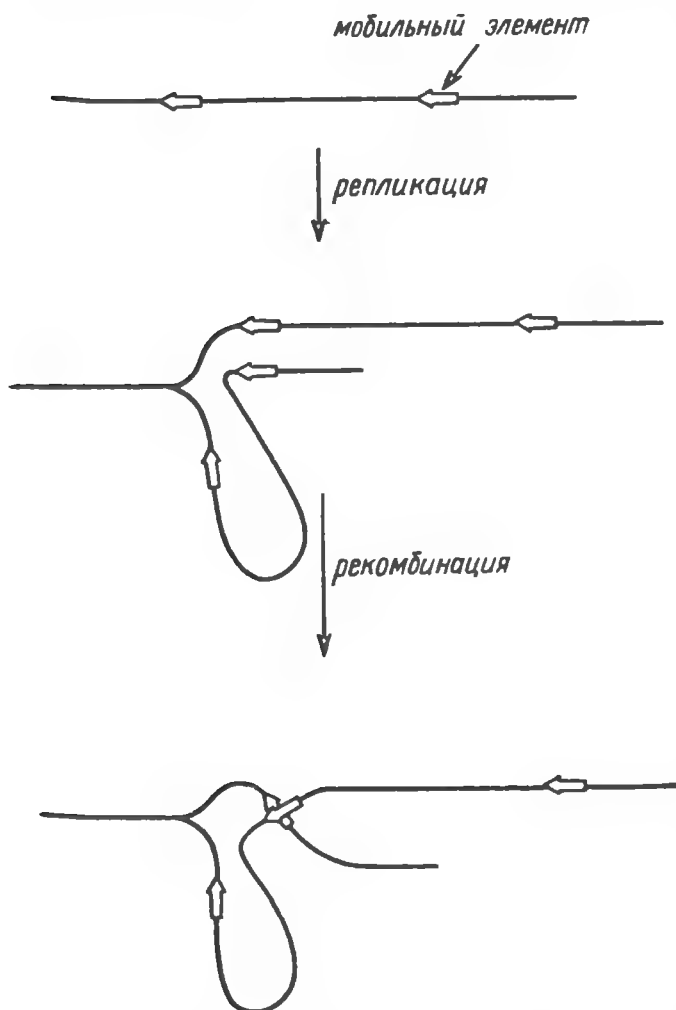
Способность многих плазмид передаваться из клетки в клетку при конъюгации также становится понятной исходя из предположения об их «эгоистичности». Действительно, такое «заражение» все новых клеток — очевидный (и, по-видимому, достаточно распространенный) способ избежать элиминации из бактериальной популяции в условиях, когда плазмиды не приносят хозяину явных селективных выгод. Цель будет достигнута, если частота переноса в среднем не меньше, чем частота спонтанной утери плазмид, или если она компенсирует несколько меньшую скорость роста содержащих плазмиду бактерий, которая, в принципе, может наблюдаться из-за необходимости реплицировать дополнительный генетический материал. В этой связи необходимо отметить, что присутствие на плазмиде транспозонов и IS-элементов может расширить ее возмож-

ности передаваться новым клеткам за счет способности образовывать коинтеграаты с другими плазмидами.

Хотя, как мы убедились, многие свойства мобильных элементов и плазмид разумно рассматривать с точки зрения концепции эгоистичной ДНК, они играют существенную позитивную роль в жизни бактерий-хозяев, даже если не приносят непосредственной пользы. Дело в том, что IS-элементы, транспозоны и плазмиды служат важнейшими факторами генетической изменчивости и эволюции бактерий.

4. ПЛАЗМИДЫ, МОБИЛЬНЫЕ ЭЛЕМЕНТЫ И ГЕНЕТИЧЕСКАЯ ИЗМЕНЧИВОСТЬ БАКТЕРИЙ

Бактериальные мобильные элементы и плазмиды могут не только переносить между клетками адаптивные гены, входящие в их состав,



но и являются важными факторами генетической изменчивости. Выше уже говорилось о способности IS-элементов и транспозонов вызывать инсерционные мутации, делеции, инверсии и транслокации генетического материала. В связи с этим интересно отметить, что хромосомы родственных видов бактерий отличаются друг от друга многочисленными перестройками именно этого типа. По-видимому, мобильные элементы сыграли существенную роль в дивергенции и видообразовании бактерий. Перечисленными выше мутациями не исчерпываются возможности мобильных элементов. Встраивание некоторых IS-элементов поблизости от молчащего гена может приводить к его активации за счет транскрипции с промотора IS-элемента, т. е. такое событие способно изменить регуляцию

Рис. 81. Механизм образования дупликаций за счет гомологичной рекомбинации по мобильным элементам — «подвижным участкам гомологии» (неравный кроссинговер)

бактериального гена. Очень важно, что мобильные элементы служат подвижными участками гомологии, гомологичная рекомбинация между которыми может приводить к дупликациям генов (рис. 81). Считается, что дупликации — один из основных путей эволюционного возникновения новых функций. Действительно, «лишняя» копия гена выходит из-под давления естественного отбора и получает возможность накапливать изменения. Чаще всего это приведет к утрате какой бы то ни было функции, но иногда может получиться ген с новыми функциями.

Одинаковые IS-элементы и транспозоны, расположенные на разных репликаонах, способны обеспечивать гомологичную рекомбинацию, приводящую к образованию коинтеграта. Именно таким путем некоторые плазмиды обратимо встраиваются в хромосому бактерий, что сразу обеспечивает добавление значительного фрагмента генетического материала (рис. 82). Плазмиды, способные встраиваться в хромосому бактерий и вырезаться оттуда, называют *эписомами*. Иногда вырезание эписомы может происходить не по той паре IS-элементов, по которой прошла интеграция. В этом случае плаزمида может захватить часть хромосомного материала, а часть своей ДНК

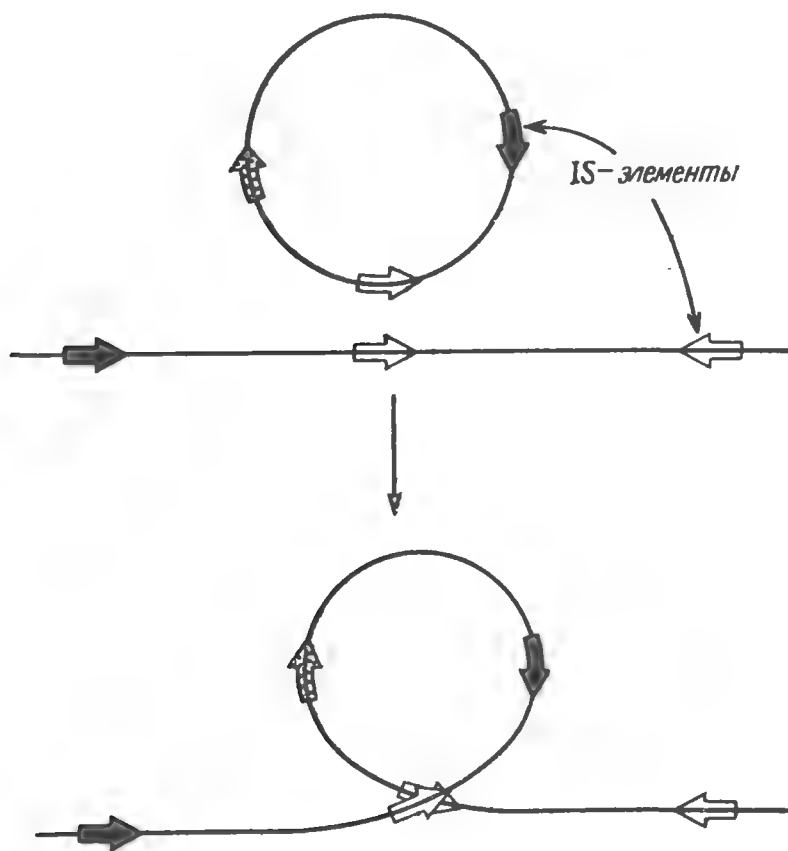


Рис. 82. Механизм интеграции некоторых плазмид (Гомологичная рекомбинация между одинаковыми IS-элементами плазмиды и хромосомы приводит к встраиванию плазмиды в хромосому. Такое встраивание обратимо, поскольку рекомбинация между той же парой IS-элементов способна вырезать интегрированную плазмиду)

может оставить в хромосоме. Все эти процессы существенно расширяют диапазон возможных изменений генетического материала.

Чрезвычайно важным является то обстоятельство, что интегрированная в хромосому конъюгативная плазмида (например, F-фактор *E. coli*) не теряет способности инициировать конъюгацию клеток и перенос ДНК из донора в реципиент. При этом ДНК плазмиды, составляющая одно целое с хромосомной ДНК, «затаскивает» в реципиент хромосому бактерии-донора. Между ДНК донора и реципиента может происходить общая рекомбинация, что приводит к обмену гомологичными генами между клетками бактериальной популяции. Этот процесс — бактериальный аналог полового размножения. Наличие механизма обмена генами очень важно для эволюции бактерий, поскольку, как и в случае полового размножения эукариот, нарушает абсолютную сцепленность генов одной хромосомы и позволяет естественному отбору находить благоприятные комбинации уже присутствующих в популяции бактерий аллельных вариантов генов.

Штаммы, в которых эписома находится в интегрированном состоянии, переносят свои гены бесплазмидным клеткам с высокой частотой и поэтому называются Hfr (от англ. high frequency of recombination — высокая частота рекомбинации). Перенос хромосомных генов при конъюгации всегда начинается с одной и той же точки — места интеграции эписомы — и имеет одинаковую направленность, определяемую ориентацией встроившейся плазмиды. Поэтому, прерывая перенос через разные промежутки времени после его начала, можно определить порядок расположения генов на бактериальной хромосоме поблизости от места интеграции плазмиды. Из-за применения этого способа картирования расстояния между генами на генетической карте *E. coli* обозначены в минутах.

Важность обмена генетическим материалом для эволюции прокариот подтверждается тем, что многие бактерии имеют другой механизм обмена генами — естественную трансформацию. В ходе этого процесса бактерии активно поглощают ДНК, оказавшуюся в среде. Если поглощенная ДНК гомологична внутриклеточной, то возможна рекомбинация между ними. Для того чтобы повысить вероятность попадания в клетку именно гомологичной ДНК, некоторые бактерии имеют систему дискриминации, узнающую определенную последовательность ДНК, часто встречающуюся у этих бактерий, но редко у других, и позволяющую транспорт в клетку лишь тех молекул ДНК, которые отмечены такой последовательностью. Проникновение в клетку произвольной ДНК из среды потенциально опасно: таким путем могли бы проникать патогенные агенты, например вирусы. Видимо, поэтому при естественной трансформации в клетку проникает лишь одна линейная цепь ДНК, а вторая в ходе транспорта деградирует. В таком виде ДНК относительно безвредна: она рекомбинирует с клеточной ДНК при наличии гомологичных участков, а при отсутствии гомологии, как правило, де-

градирует. Одноцепочечная ДНК служит очень хорошим субстратом для ферментов рекомбинации, которые охотно «вставляют» ее в гомологичный дуплекс (см. рис. 58 и 59). Поэтому доля рекомбинационных событий, приходящихся на единицу длины ДНК, заметно выше при естественной трансформации, чем при конъюгативном переносе, когда рекомбинация происходит между двумя дуплексами ДНК.

ЛИТЕРАТУРА

П. Брода. Плазмиды.— М.: Мир, 1982.

Р. Б. Хесин. Непостоянство генома.— М: Наука, 1984.

Г. Стент, Р. Кэлиндар. Молекулярная генетика.— М: Мир, 1981.

Б. Льюин. Гены.— М.: Мир, 1987.

Глава VI

РЕСТРИКЦИЯ И МОДИФИКАЦИЯ ДНК

Широкое распространение обмена ДНК между бактериями ставит перед ними задачу сохранения собственного генома. Далеко не всегда проникшая в клетку чужеродная ДНК способна оказаться полезной. Более того, посторонний генетический материал может быть губительным для клетки, особенно если принадлежит бактериальному вирусу, бактериофагу. Для того чтобы бороться с чужеродной ДНК, нужно уметь отличать свою ДНК от чужой. Бактерии достигают этого тем, что «метят» свою ДНК с помощью специального модифицирующего фермента. Практически все виды бактерий имеют *метилазы*, модифицирующие аденин или цитозин в определенной, характерной для данного вида последовательности ДНК. Другой специальный фермент, *эндонуклеаза рестрикции* (рестриктаза), узнает ту же последовательность и разрезает ее, если она не модифицирована, т. е. попала в клетку извне. Таким путем бактерии ограничивают возможности попадания в них постороннего генетического материала.

1. РЕСТРИКЦИЯ БАКТЕРИОФАГОВ

Явление рестрикции состоит в том, что бактериофаг, размноженный на одном штамме бактерий, часто очень плохо размножается на другом штамме. Те редкие фаги, которые все-таки смогли размножиться на «чужом» для них штамме, оказываются адаптированными к нему, или модифицированными, т. е. следующий раунд размножения проходит с полной эффективностью. Одновременно эти модифицированные фаги теряют способность эффективно размножаться на том штамме, на котором были выращены исходно (табл. 6).

Таблица 6. Эффективность посева фага λ на разных штаммах

Штамм, на котором был размножен фаг	Эффективность посева	
	К	В
К	1	10^{-4}
В	10^{-4}	1

Сейчас известно, что рестрикция связана с расщеплением фаговой ДНК эндонуклеазой, специфически узнающей определенную последовательность нуклеотидов, а модификация состоит в метилировании той же последовательности. Различают системы рестрикции и модификации I, II и III типа.

2. СИСТЕМЫ РЕСТРИКЦИИ

Системы рестрикции II типа

Системы рестрикции II типа обнаружены у очень многих бактерий. Эти системы состоят из двух отдельных ферментов, рестриктазы и метилазы, узнающих одну и ту же последовательность ДНК — сайт рестрикции. Если сайт рестрикции не метилирован, то рестриктаза вносит в него двуцепочечный разрыв. ДНК не подвергается рестрикции, если хотя бы одна цепь метилирована. Такие свойства предохраняют собственную ДНК бактерий от рестрикции: собственная ДНК либо полностью метилирована по всем сайтам рестрикции, либо, после репликации, полуметилирована. На полуметилированные сайты рестрикции действует метилаза и метилирует их полностью. Как и прочие клеточные метилазы, метилазы системы рестрикции-модификации в качестве донора метильных групп используют S-аденозилметионин. В табл. 7 для примера приведены данные о некоторых рестриктазах и метилазах II типа.

Сайт рестрикции рестриктаз II типа, как правило, симметричен: обе цепи ДНК имеют одинаковую последовательность, если читать ее в одном и том же направлении (например, от 5'-к 3'-концу ДНК) для каждой цепи (табл. 7). Это естественно, поскольку рестриктазы вносят разрывы в обе цепи ДНК и метилазы тоже метилируют обе цепи.

Рестриктазы II типа очень широко используются в методах генной инженерии для физического картирования ДНК и для выделения участков ДНК в составе того или иного рестрикционного фрагмента. Поэтому в течение ряда лет велся широкий поиск рестриктаз II типа с разной специфичностью (разными сайтами рестрикции). В результате сейчас известно более 350 рестриктаз разных бакте-

Таблица 7. Некоторые рестриктазы второго типа

Фермент рестрикции ¹	Сайт узнавания ²	Источник
Bst I	5' — GGATCC — 3' CCTAGG	<i>Bacillus stearothermophilus</i>
Hind III	AAGC TT TTCG AA	<i>Haemophilus influenzae</i> Rd
Alu I	AGCT TCGA	<i>Arthrobacter luteus</i>
Pst I	CTGCAG GACGTC	<i>Providencia stuartii</i>

¹ Соответствующие метилазы называются M-Bst I, M-Alu I и т. д.
² Стрелками отмечены фосфодиэфирные связи, разрываемые рестриктазами, звездочками — основания, модифицируемые соответствующей метилазой.

рий, узнающих более 80 разных сайтов (рестриктазы из разных источников, узнающие один сайт, называют *изошизомерами*). Лишь для считанных ферментов действительно показано, что они участвуют в описанном выше явлении рестрикции.

Системы рестрикции I и III типа

Системы рестрикции I типа были открыты раньше, но оказались значительно сложнее систем II типа. Рестриктазы I типа известны у штаммов *E. coli* В и К (и называются соответственно рестриктазами EcoB и EcoK) и у разных штаммов *S. typhimurium*. В этих системах и рестрикцию, и модификацию осуществляет один фермент, состоящий из трех субъединиц, кодируемых у *E. coli* генами *hsdM*, *hsdS* и *hsdR*. Субъединица, кодируемая геном *hsdS*, отвечает за специфичность фермента, она ответственна за узнавание сайта рестрикции. Субъединица, кодируемая геном *hsdR*, ответственна за рестрикцию, а субъединица, кодируемая геном *hsdM* — за модификацию. Фермент узнает последовательность несколько более сложную, чем в случае рестриктаз II типа, например для EcoB: 5' = TG $\ddot{A}$ N $\ddot{A}$ TGCT = 3'. Если оба отмеченных звездочкой аденина этой последовательности метилированы, то рестриктаза ничего с последовательностью не делает. Если метилирован только один аденин (например, после репликации или репарации), то метилаза метилирует второй, используя S-аденозилметионин в качестве донора метильной группы. Если же сайт не метилирован, то ДНК, в составе

которой он находится, распознается как чужеродная и происходит рестрикция. Для рестрикции также необходим S-аденозилметионин, но лишь в качестве аллостерического эффектора. Подобная регуляция страхует клетку от рестрикции собственной ДНК в условиях недостатка донора метильных групп, когда не может эффективно происходить модификация ДНК. И для рестрикции, и для модификации необходим АТР, причем при модификации АТР играет лишь аллостерическую роль, а для рестрикции необходим гидролиз АТР. Дело в том, что по малопонятным причинам рестриктазы I типа вносят разрыв в ДНК на огромном расстоянии от сайта узнавания — не менее 1000 п. о. Энергия гидролиза АТР затрачивается на перемещение фермента от сайта узнавания до участка разрезания ДНК, точнее фермент, оставаясь связанным с сайтом узнавания, «протягивает» через себя ДНК до тех пор, пока не внесет разрыв.

Подобная сложность систем рестрикции I типа привела к предположению о том, что эти ферменты играют в клетках какую-то роль помимо рестрикции. Однако другие функции рестриктаз пока не обнаружены.

Бактерии, мутантные по гену *hsdR*, не способны к рестрикции, не могут модифицировать ДНК (r^-m^-). Мутация в гене *hsdS* приводит к тому, что фермент не может узнать свой сайт, поэтому в таких мутантах отсутствует и рестрикция, и модификация (r^-m^-). Мутанты по гену *hsdM* нежизнеспособны, так как дефект метилазы при наличии активной рестриктазы приводит к деградации собственной ДНК. *hsdR* и S-мутанты *E. coli* широко используют в генетической инженерии, поскольку в такие клетки можно вводить протяженную чужеродную, т. е. немодифицированную, ДНК, которая при этом не подвергается рестрикции.

Рестриктазы III типа являются промежуточными между ферментами I и II типов. Подобно ферментам I типа, в состав единого фермента III типа входят и рестриктаза, и метилаза. Рестриктазы III типа вносят разрыв не в сайт узнавания, но в отличие от рестриктаз I типа на небольшом фиксированном расстоянии от него.

3. БОРЬБА БАКТЕРИОФАГОВ С СИСТЕМАМИ РЕСТРИКЦИИ ХОЗЯИНА

Защита собственного генетического материала с помощью системы рестрикции не всегда эффективна, о чем свидетельствует само существование бактериальных вирусов — бактериофагов. Оказывается, бактериофаги выработали разнообразные тактики борьбы с рестрикцией. Например, для сравнительно небольших фагов известны случаи, когда жесткий отбор на преодоление рестрикции привел к полному отсутствию сайтов узнавания рестриктазы хозяина на фаговой ДНК. Другой способ борьбы с рестрикцией используют некоторые крупные бактериофаги. В состав их ДНК входит необычное основание, например в ДНК Т-четных фагов *E. coli*

(Т 2, Т4 и Т6) вместо цитозина входит оксиметилцитозин, большинство остатков которого к тому же еще гликозилированы. ДНК, модифицированную подобным образом, не разрезают почти все известные рестриктазы. Такой способ борьбы с рестрикцией могут позволить себе лишь крупные фаги, поскольку для этого фаг должен кодировать фактически новый метаболический путь синтеза необычных нуклеотидов, а также новые ферменты синтеза ДНК, адаптированные к необычным свойствам фаговой матрицы и к использованию необычного субстрата. Зато столь радикальная модификация ДНК позволяет фагам, использующим эту тактику, не только защищаться от хозяйских рестриктаз, но и пойти дальше: они кодируют нуклеазы, деградирующие немодифицированную ДНК хозяина, но не действующие на фаговую ДНК. Другие фаги используют еще ряд способов борьбы с системой рестрикции хозяина.

У эукариотических клеток, по-видимому, отсутствует аналог бактериальной системы рестрикции-модификации. Хотя у эукариот имеется метилированная ДНК и по крайней мере у некоторых — сайт-специфические эндонуклеазы (см. гл. IV), они не образуют единую защитную систему. Таким образом, защита эукариотических клеток от вирусов основана на совершенно других механизмах.

ЛИТЕРАТУРА

- Г. Стент, Р. Кэлиндар. Молекулярная генетика. — М.: Мир, 1981.
Г. Стент. Молекулярная биология вирусов бактерий. — М.: Мир, 1965.
Дж. Уотсон, Дж. Туз, Д. Курц. Рекомбинантные ДНК. — М.: Мир, 1986.

Глава VII

ТРАНСКРИПЦИЯ

Транскрипция является первой стадией реализации (считывания) генетической информации, на которой нуклеотидная последовательность ДНК копируется в виде нуклеотидной последовательности РНК. В основе механизма копирования при транскрипции лежит тот же структурный принцип комплементарного спаривания оснований, что и при репликации. Транскрипция осуществляется ферментами РНК-полимеразами, синтезирующими РНК на ДНК-матрице из рибонуклеозидтрифосфатов.

Синтез молекул РНК начинается в определенных местах ДНК, называемых *промоторами*, и завершается в *терминаторах* (рис. 83). Участок ДНК, ограниченный промотором и терминатором, представляет собой единицу транскрипции — *транскриптон*. В пределах каждого транскриптона копируется только одна из двух нитей ДНК, которая называется *значащей* или *матричной*. Во всех транскриптонах, считываемых в одном направлении, значащей является

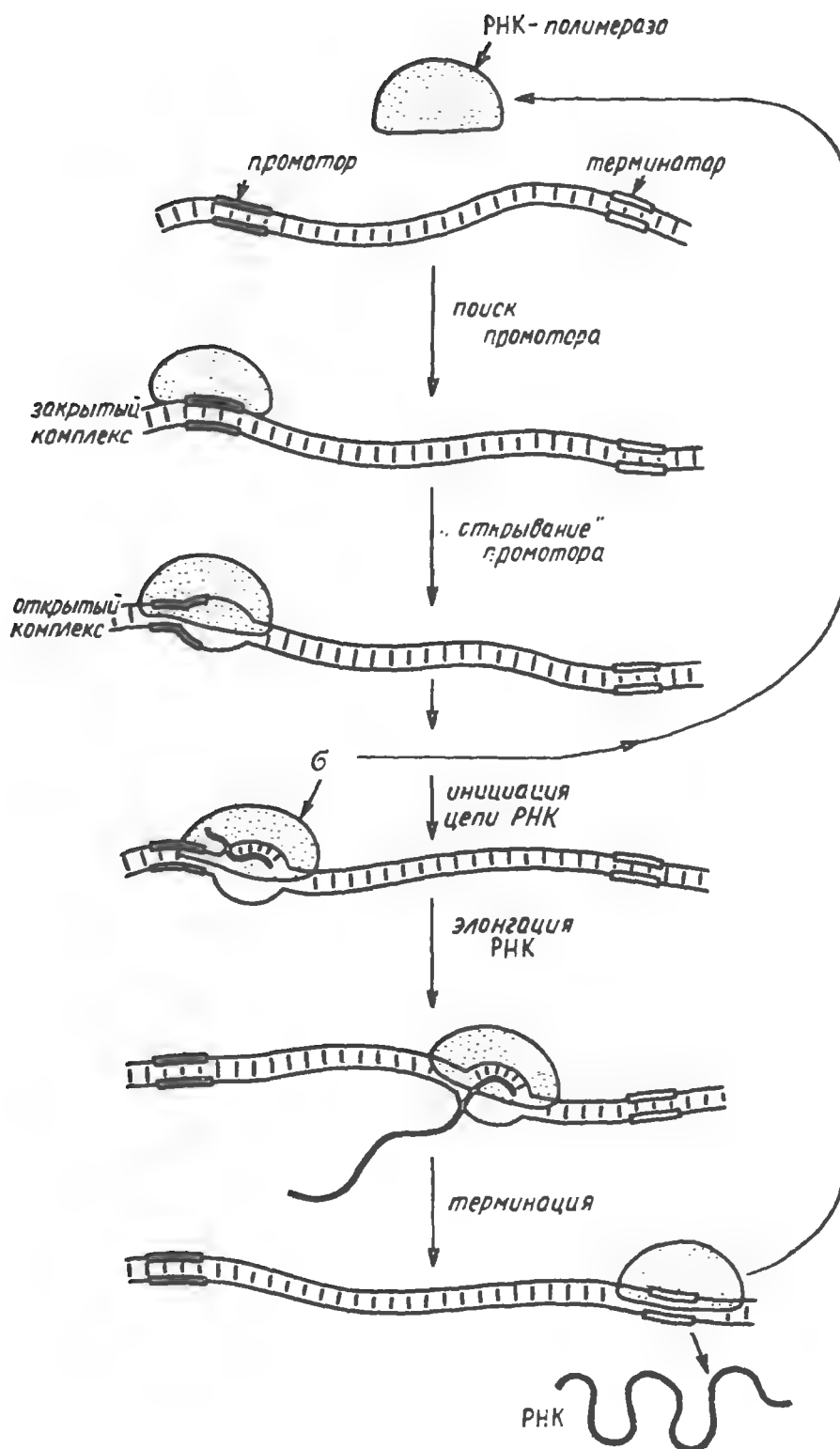


Рис. 83. Цикл транскрипции

одна нить ДНК; в транскриптонах, считываемых в противоположном направлении, значащей является другая нить ДНК. Соседние транскриптоны могут быть отделены друг от друга нетранскрибируемыми участками ДНК, а могут и перекрываться, в частности так,

что в пределах участка перекрывания матричными оказываются обе нити. Разбиение ДНК на множество транскриптонов обеспечивает возможность независимого считывания разных генов, их индивидуального включения и выключения.

У эукариот в состав транскриптона, как правило, входит только один ген. Транскриптоны прокариот чаще называют *оперонами*; многие из них содержат по несколько генов, обычно функционально связанных, например, кодирующих ферменты синтеза той или иной аминокислоты. Существуют опероны, содержащие гены, не кодирующие белков (гены рибосомных РНК, тРНК и др.). Описаны смешанные опероны, включающие гены тРНК и белков.

1. РНК-ПОЛИМЕРАЗЫ

У бактерий молекула РНК-полимеразы состоит из двух компонентов: минимальной РНК-полимеразы, или «кор»-фермента (от англ. core — сердцевина), содержащего все каталитические центры, участвующие в синтезе РНК, и σ -субъединицы, необходимой для правильного присоединения фермента к промотору и отделяющейся от РНК-полимеразы после начала синтеза РНК. Минимальная РНК-полимераза состоит из четырех субъединиц: двух идентичных α -субъединиц и неидентичных β - и β' -субъединиц. В β - и β' -субъединицах имеется по одному атому Zn. У *E. coli* α -субъединица состоит из 329 аминокислотных остатков ($M_r=36,51$ кД), β — из 1342 (150,62 кДа), а β' — из 1407 (155,16 кДа). Размеры этих субъединиц у других бактерий примерно такие же. Сильно вытянутые β - и β' -субъединицы имеют длину около 24,0 нм (что соответствует 70 п. н. в ДНК) и ширину примерно 40 нм. Эти субъединицы контактируют с ДНК в комплексе РНК-полимеразы с промотором и несут в себе каталитические центры.

β -Субъединица формирует центр связывания антибиотиков рифампицина и стрептолидигина, являющихся специфическими ингибиторами бактериальных РНК-полимераз. Функция α -субъединицы не известна.

Транскрипцию генов рибосомных РНК, тРНК и большинства генов, кодирующих белки, обеспечивают молекулы РНК-полимеразы, содержащие «главную» σ -субъединицу (молекулярная масса у *E. coli* 70 кД, у *Bac. subtilis* — 43 кД). На несколько тысяч молекул РНК-полимеразы, имеющих в бактериальной клетке, приходится примерно тысяча молекул главной σ -субъединицы. В меньших количествах имеются «минорные» σ -субъединицы, используемые для транскрипции ограниченного числа генов (см. раздел 3 этой главы). Набор минорных σ -субъединиц у разных бактерий неодинаков. По размеру они меньше главной σ -субъединицы. Сравнение нуклеотидных последовательностей генов разных σ -субъединиц свидетельствует о том, что все они произошли от одного предкового гена.

По аминокислотной последовательности субъединицы РНК-полимераз далеких видов бактерий существенно различаются, но по пространственной структуре, по-видимому, весьма сходны, так как в пробирке удастся собирать активные гибридные молекулы, часть субъединиц которых взяты от одного вида бактерий, а часть — от другого.

В ядрах эукариот обнаружены три специализированные формы РНК-полимеразы. РНК-полимераза I (или А по другой номенклатуре) осуществляет в ядрышке синтез 18 S и 23 S рРНК. РНК-полимераза II (или В) — синтез информационных РНК. РНК-полимераза III (или С) — синтез тРНК и некоторых других низкомолекулярных РНК. РНК-полимеразы эукариот нечувствительны к ингибиторам бактериальных РНК-полимераз — рифампицину и стрептолидигину. Специфическим ингибитором РНК-полимеразы II является токсин бледной поганки α -аманитин. Каждая форма РНК-полимеразы состоит из двух больших субъединиц с $M_r \simeq 120 \div 220$ кД и от 5 до 13 малых с $M_r = 10 \div 100$ кД. Несколько малых субъединиц являются общими для разных форм РНК-полимераз. Большие субъединицы, по-видимому, выполняют те же функции, что β - и β' -субъединицы бактериальных РНК-полимераз. О функциях малых субъединиц пока ничего не известно.

В самых больших субъединицах эукариотических РНК-полимераз обнаружено несколько участков, которые по аминокислотной последовательности у всех трех форм сходны между собой и с β' -субъединицей РНК-полимеразы *E. coli*. В следующих за ними по размеру субъединицах эукариотических РНК-полимераз обнаружено сходство в аминокислотной последовательности с β -субъединицей РНК-полимеразы *E. coli*. Эти данные свидетельствуют о том, что на заре эволюции эукариот у них имелась одна форма РНК-полимеразы, а разные формы возникли за счет умножения предковых генов (общих для про- и эукариот) и последующего расхождения их нуклеотидных последовательностей в результате множества мутаций.

Особые РНК-полимеразы обеспечивают транскрипцию клеточных органелл эукариот — хлоропластов и митохондрий. В составе хлоропластной ДНК обнаружены гены, гомологичные генам, кодирующим α -, β - и β' -субъединицы РНК-полимеразы *E. coli*. Это, а также сходство нуклеотидной последовательности промоторов бактерий и хлоропластов свидетельствует о том, что РНК-полимераза хлоропластов должна быть сходна с РНК-полимеразой бактерий. РНК-полимеразы митохондрий состоят, по-видимому, всего из одной субъединицы, подобно РНК-полимеразам, кодируемым некоторыми бактериофагами, такими, как Т3 и Т7. РНК-полимераза митохондрий дрожжей сходна с РНК-полимеразами этих фагов по аминокислотной последовательности. Ген, кодирующий митохондриальную РНК-полимеразу, располагается в ядре.

Десять лет назад было выделено совершенно новое царство кле-

точных организмов — архебактерии. Архебактерии — это прокариотические микроорганизмы, многие из которых обитают в экстремальных условиях: при высоких концентрациях солей или при высоких температурах, приближающихся к 100 °С. Анализ нуклеотидных последовательностей их рибосомных РНК показал, что архебактерии примерно одинаково далеко отстоят по своей эволюционной истории и от эукариот, и от обычных бактерий. У архебактерий не обнаружено множества форм РНК-полимераз, но по своему субъединичному составу их РНК-полимеразы похожи на РНК-полимеразы эукариот: в их состав входит 8—10 субъединиц. Весьма четко прослеживается гомология больших субъединиц РНК-полимераз различных архебактерий с соответствующими субъединицами РНК-полимераз эукариот.

2. ЦИКЛ ТРАНСКРИПЦИИ

Цикл транскрипции можно разделить на четыре основные стадии, каждая из которых в свою очередь состоит из многих элементарных этапов: 1) связывание с ДНК; 2) инициация цепи РНК; 3) рост (элонгация) цепи РНК; 4) терминация цепи РНК.

Этот цикл у бактерий удается целиком осуществить в простой бесклеточной системе, состоящей из ДНК-матрицы и очищенной РНК-полимеразы, без каких бы то ни было дополнительных факторов. Это не значит, что РНК-полимераза является единственным белком, участвующим в транскрипции. В ней могут участвовать и разнообразные регуляторные белки. Однако роль их вспомогательная: они мешают или помогают РНК-полимеразе на тех или иных стадиях цикла транскрипции, которые она осуществляет и в их отсутствие. Поэтому изучение цикла транскрипции изолированной бактериальной РНК-полимеразой позволяет понять не только ферментативные механизмы синтеза молекулы РНК, но, что еще важнее, дает ключ к пониманию механизмов регуляции транскрипции.

У эукариот транскрипция изучена хуже, чем у бактерий. Это связано, в частности, с тем, что очищенные РНК-полимеразы эукариот не способны осуществить полный цикл транскрипции. Для этого всегда нужны дополнительные белковые факторы, лишь часть из которых получена в очищенном виде. Пока эукариотический цикл транскрипции удастся осуществить лишь в частично очищенных клеточных экстрактах, содержащих недостаточно охарактеризованные компоненты. Тем не менее можно считать, что в основных чертах цикл транскрипции эукариот и бактерий сходен.

В этом разделе этот цикл будет описан на примере хорошо изученного цикла для РНК-полимеразы *E. coli* (см. рис. 83).

Цикл транскрипции начинается с *присоединения РНК-полимеразы к промотору* — строго определенному участку ДНК, с которого начинается синтез РНК. Механизм поиска промоторов изучен недостаточно; предполагается, что молекулы РНК-полимеразы при-

соединяются к случайным местам двойной спирали ДНК, некоторое время перемещаются по ней, отсоединяются и снова присоединяются к ДНК до тех пор, пока не окажутся на промоторах.

Оказавшись на промоторе, РНК-полимераза образует с ним так называемый закрытый промоторный комплекс, в котором ДНК сохраняет двуспиральную структуру. В закрытом комплексе РНК-полимераза еще не способна к синтезу РНК. Этот комплекс нестабилен и легко диссоциирует при повышении ионной силы.

Закрытый комплекс может обратимо превращаться в открытый комплекс, в котором РНК-полимераза расплетает примерно один виток двойной спирали ДНК в районе *стартовой точки* — нуклеотида, с которого начинается комплементарное копирование матрицы. В открытом комплексе связь РНК-полимеразы с промотором становится значительно более прочной, чем в закрытом.

При низких температурах равновесие сильно смещено в сторону закрытого комплекса, при температурах, близких к физиологическим, — в сторону открытого. Переход между этими крайними случаями происходит в узком температурном интервале, что, по-видимому, отражает кооперативный характер плавления ДНК промотора.

Следующая стадия, *инициация*, требует наличия субстратов РНК-полимеразы, нуклеозидтрифосфатов и заключается в образовании первых нескольких звеньев цепи РНК. Первый нуклеотид входит в состав цепи, сохраняя свою трифосфатную группу, а последующие присоединяются к 3'-ОН-группе предыдущего с освобождением пирофосфата. На стадии инициации РНК-продукт связан с матрицей и РНК-полимеразой непрочно и с высокой вероятностью может освобождаться из комплекса. В этом случае РНК-полимераза, не покидая промотора, снова иницирует РНК. Такой синтез ди-, три- и более длинных олигонуклеотидов называют абортивной инициацией в противоположность продуктивной (т. е. завершающейся образованием полноценного РНК-продукта) инициации. Когда РНК-продукт достигает критической длины (от 3 до 9 нуклеотидов на разных промоторах), абортивная инициация полностью прекращается, транскрибирующий комплекс стабилизируется и уже не распадается до тех пор, пока синтез молекулы РНК не будет доведен до конца. Примерно в этот же момент, который считается концом инициации и началом элонгации, от РНК-полимеразы отделяется σ -субъединица.

Эффективность инициации на разных промоторах, их «сила», существенно различается: если с некоторых промоторов иницируется всего одна-две молекулы РНК за период деления клетки, то с других (например, с промоторов генов рибосомных РНК) инициация происходит раз в одну-две секунды. Частота, с которой иницируется транскрипция при насыщающей концентрации субстратов, зависит главным образом от равновесной константы образования закрытых промоторных комплексов и константы скорости превра-

щения закрытого комплекса в открытый. Для самых сильных промоторов обычно характерны высокие значения обеих констант (высокое сродство РНК-полимеразы к промотору и быстрый переход промоторного комплекса в активное открытое состояние). Для слабых промоторов характерны низкие значения этих величин. Слабость промоторов с низким сродством к РНК-полимеразе особенно заметна при низких концентрациях РНК-полимеразы и может быть скомпенсирована при ее высоких концентрациях. Промотор с низким сродством к РНК-полимеразе может быть достаточно сильным, если ему присуща высокая скорость перехода в открытое состояние. Сила большинства промоторов увеличивается с увеличением степени отрицательной сверхспирализации ДНК. Это объясняется тем, что отрицательная сверхспирализация облегчает расплетание ДНК и тем самым переход в открытый промоторный комплекс. Существуют, однако, промоторы, сила которых не зависит или даже уменьшается с увеличением степени сверхспирализации. Этому эффекту объяснения пока не найдено.

На стадии *элонгации* в ДНК расплетено примерно 18 н. п. Примерно 12 нуклеотидов матричной нити ДНК образует гибридную спираль с растущим концом цепи РНК (рис. 84). По мере движения РНК-полимеразы по матрице впереди нее происходит расплетание, а позади — восстановление двойной спирали ДНК. Одновременно освобождается очередное звено растущей цепи РНК из комплекса с матрицей и РНК-полимеразой. Эти перемещения должны сопровождаться относительным вращением РНК-полимеразы и ДНК. Трудно себе представить, как это может происходить в клетке, особенно при транскрипции хроматина. Поэтому не исключено, что для предотвращения такого вращениядвигающуюся по ДНК РНК-полимеразу сопровождают топоизомеразы.

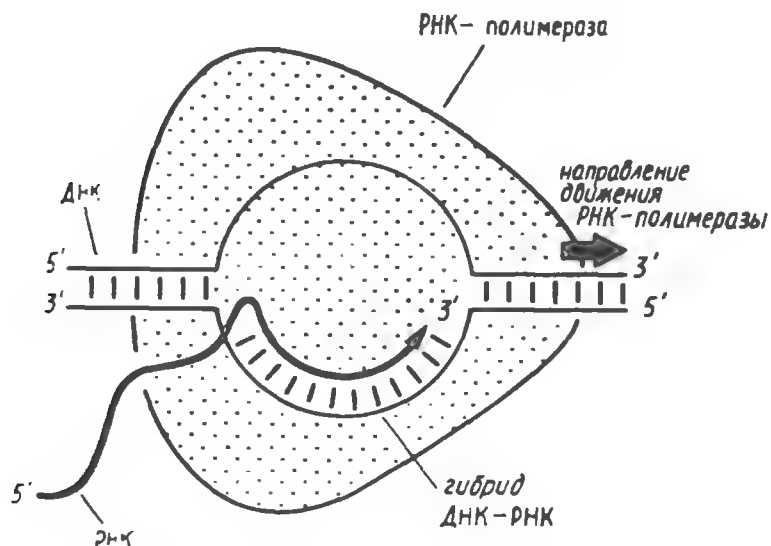


Рис. 84. РНК-полимераза, транскрибирующая ДНК

В процессе удлинения РНК РНК-полимераза движется по ДНК с непостоянной скоростью. Особенно четко это видно при низких концентрациях субстратов. В некоторых участках матрицы длительные задержки в продвижении РНК-полимеразы, так называемые *паузы*, наблюдаются даже при оптимальных концентрациях субстратов. Завершается синтез РНК в строго определенных участках матрицы — терминаторах, где происходит отделение от ДНК и готовой РНК, и минимальной РНК-полимеразы, которая, объединившись со свободной σ -субъединицей, может вступить в следующий цикл транскрипции.

Нуклеотидные последовательности промоторов прокариот

Основной элемент промотора — место связывания РНК-полимеразы, которое она занимает перед началом синтеза РНК. В состав промоторов могут входить также участки связывания белков-регуляторов. Размер участка связывания РНК-полимеразы соответствует ее длине и составляет примерно 70 п. н. Располагается этот участок относительно стартовой точки несимметрично: по ходу транскрипции его граница отстоит от стартовой точки на 20 п. н., а против хода — примерно на 50 п. н. (рис. 85).

При сравнении нуклеотидной последовательности большого числа промоторов *E. coli*, узнаваемых РНК-полимеразой, содержащей главную σ -субъединицу, оказалось, что одинаковых среди них нет. Сходство между ними обнаружилось в основном в двух участках длиной по 6 п. н., центры которых располагаются в районах «—10» и «—35» п. н. (нумерация нуклеотидов промотора ведется от стартовой точки, которой приписывается номер +1; рис. 85). Некоторое сходство наблюдается также в районе стартовой точки.

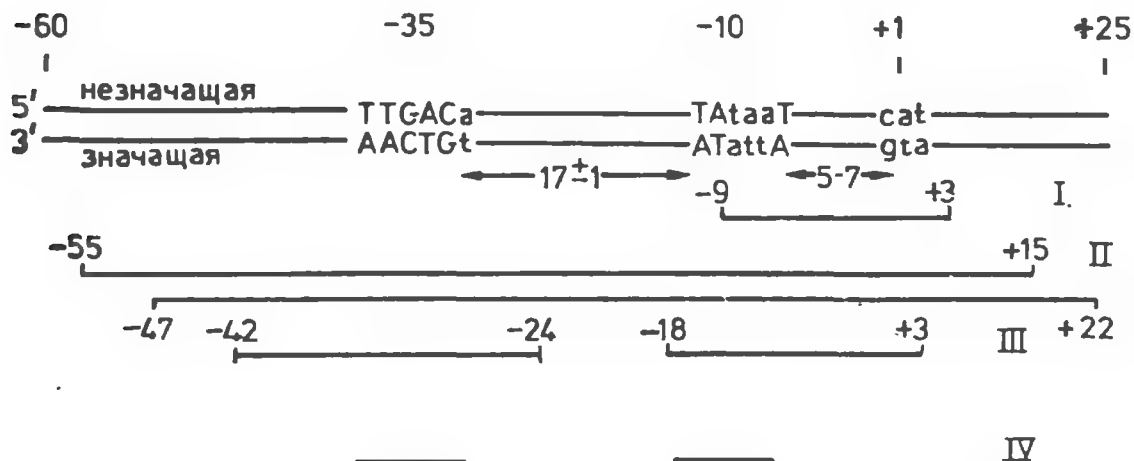


Рис. 85. Структура промотора:

I — участок, расплетенный в открытом комплексе; II — участок, защищенный от ДНКазы I; III — участки расположения основных «контактов» — мест изменения модификации диметилсульфатом; IV — участки локализации промоторных мутаций

На основании сравнения последовательностей разных промоторов выведена «каноническая» последовательность промотора, в которой представлены наиболее часто встречающиеся в каждом положении нуклеотиды. Каноническая последовательность участка «—10» — ТАТААТ (эта последовательность называется также *блоком Прибнова*), участки «—35» — TTGACA (при рассмотрении промоторов обычно приводят последовательность только той нити ДНК, которая в транскрибируемой части совпадает с последовательностью РНК, т. е. является *незначащей*). Каноническая последовательность промотора несимметрична, что отражает его функциональную несимметричность. Действительно, промотор определяет не только место начала транскрипции, но и ее направление. Среди природных промоторов пока не обнаружено ни одного с канонической последовательностью, но искусственно сконструированный промотор с канонической последовательностью отличается очень высокой эффективностью (этот результат не был заранее очевиден: усредненная последовательность вполне могла бы обладать «средними» свойствами). О том, что каноническая последовательность является наиболее эффективной, свидетельствуют и результаты многочисленных данных по мутационным изменениям последовательности промоторов: изменения, приближающие последовательность промотора к канонической, как правило, увеличивают его силу, тогда как изменения, уменьшающие его сходство с канонической, — уменьшают его силу. Изменения нуклеотидной последовательности вне участков «—10» и «—35» обычно слабо сказываются на силе промотора. Знание этих закономерностей, однако, еще не позволяет надежно предсказывать силу промоторов и находить промоторы, рассматривая последовательность ДНК, хотя РНК-полимераза делает это очень быстро.

У разных промоторов расстояния между участками «—10» и «—35» могут быть не одинаковыми, у большинства промоторов оно составляет 16—18 п. н.; у небольшого числа промоторов вариации составляют 15—20 п. н. Оптимальное расстояние 17 п. н.: мутации, изменяющие это расстояние как в сторону увеличения, так и уменьшения, ослабляют промотор.

Стартовая точка транскрипции отстоит у разных промоторов на 6—9 п. н. от последовательности Прибнова (рис. 85). Начальным нуклеотидом РНК чаще всего является диктуемый матрицей А или G. Предпочтение пуринов, однако, неабсолютно, и есть несколько промоторов, на которых транскрипты начинаются с С или U, несмотря на наличие на расстоянии одного-двух нуклеотидов потенциального пуринового начала. У некоторых промоторов стартовая точка не строго фиксирована: инициация происходит с двух-трех соседних нуклеотидов матрицы.

Описанная выше организация промоторов свойственна не только *E. coli*. Так, например, промоторы, узнаваемые с помощью главной σ -субъединицы *Bac. subtilis*, по своей нуклеотидной последова-

тельности проявляют большое сходство с каноническим промотором *E. coli*. Поэтому не удивительно, что они хорошо используются РНК-полимеразой *E. coli*. Любопытно, что среди промоторов *Bac. subtilis* реже, чем в случае промоторов *E. coli*, встречаются значительные отклонения от канонической последовательности. Создается впечатление, что РНК-полимераза *Bac. subtilis* более ограничена в отношении выбора последовательностей нуклеотидов промоторов, чем РНК-полимераза *E. coli*. В соответствии с этим многие промоторы *E. coli* плохо используются РНК-полимеразой *Bac. subtilis*.

Промоторы, используемые РНК-полимеразами, содержащими минорные σ -субъединицы, заметно отличаются по нуклеотидной последовательности от промоторов, используемых РНК-полимеразой, содержащей главную σ -субъединицу. Для каждого типа σ -субъединицы характерна своя каноническая последовательность участков, аналогичных участкам «—35» и «—10».

3. РЕГУЛЯЦИЯ АКТИВНОСТИ ПРОМОТОРОВ ПРОКАРИОТ

Активность многих промоторов регулируется с помощью особых белков-регуляторов, которые присоединяются к определенным участкам ДНК и либо мешают, либо помогают РНК-полимеразе инициировать синтез РНК. В первом случае говорят о *негативной*, во втором — о *позитивной* регуляции активности промотора.

Белки, осуществляющие негативную регуляцию, называются *репрессорами*. Места их связывания на ДНК называются *операторами*. Способность многих репрессоров связываться со своими операторами зависит от низкомолекулярных лигандов — *эффекторов*. Эффекторы, снижающие сродство репрессора к оператору, называются *индукторами*. В отсутствие индуктора репрессор связывается с оператором и мешает РНК-полимеразе начинать синтез РНК с промотора (промотор репрессирован). В комплексе с индуктором репрессор теряет способность связываться с оператором, в результате чего промотор активируется (индуцируется). Другие репрессоры, наоборот, могут связываться с оператором только в комплексе с эффектором (который в этом случае называется *корепрессором*). В присутствии корепрессора промотор неактивен (репрессирован), в отсутствие корепрессора активируется (дерепрессируется).

Белки, осуществляющие позитивную регуляцию, называются *активаторами*. Ряд белков-регуляторов могут выступать как в роли репрессора, так и в роли активатора.

Для понимания механизмов взаимодействия РНК-полимеразы с промоторами и с белками регуляторами важно знать пространственную структуру их комплексов с ДНК. К сожалению, в настоящее время почти ничего не известно о деталях пространственной структуры РНК-полимеразы и, в частности, о структуре ее участков, взаимодействующих с ДНК. Приблизительное

представление о структуре комплексов белков с ДНК дает анализ расположения на ДНК химических групп, контактирующих с белками. Метод основан на том, что присоединение белка к ДНК защищает ее в местах контактов от химических модификаций. В другом варианте следят за тем, какие химические модификации мешают белку присоединяться к ДНК. Используют различные химические реагенты, вызывающие в конечном итоге разрыв цепи ДНК в месте модификации (принцип регистрации мест разрывов объясняется на рис. 86). В зависимости от используемого реагента можно идентифицировать контакты с фосфатными группами ДНК, метильными группами гуанина и аденина, выступающими в большую и малую бороздки, локализовать расплавленные участки ДНК. Если нанести на пространственную модель ДНК зарегистрированные таким образом контакты с белком, они очертят площадку, на которую садится белок, его «отпечаток» на ДНК (рис. 86). В случае РНК-полимеразы основные контакты группируются в районах «-35» и «-10», что еще раз подчеркивает их функциональную значимость в связы-

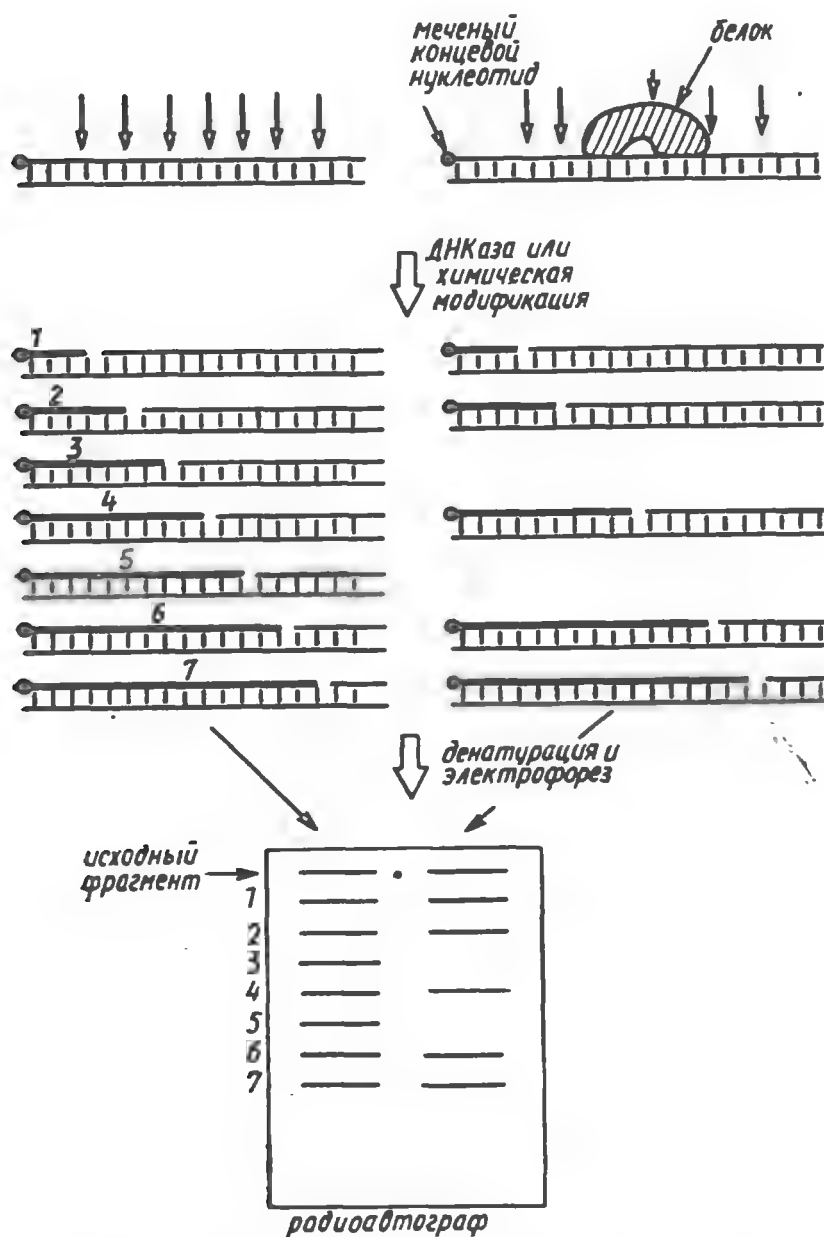


Рис. 86. Локализация контактов белка с ДНК

вании РНК-полимеразы с промотором. Рассмотрение отпечатка РНК-полимеразы на промоторе показывает, что она покрывает всю поверхность ДНК на месте своего присоединения, оставляя на другой стороне ДНК свободное место для присоединения других белков как в районе «—35», так и в районе стартовой точки.

Простейший механизм репрессии заключается в стерическом блокировании репрессором присоединения РНК-полимеразы к промотору. Такой механизм имеет место в тех промоторах, в которых участок связывания репрессора перекрывается с участком связывания РНК-полимеразы. Простейший механизм активации заключается в том, что белок-активатор присоединяется к промотору рядом с РНК-полимеразой и за счет непосредственного контакта с ней облегчает образование открытого промоторного комплекса. Дискуссионными являются механизмы действия тех белков-регуляторов, которые присоединяются к ДНК на значительном расстоянии от РНК-полимеразы. Ниже рассмотрено несколько наиболее хорошо изученных примеров, иллюстрирующих различные принципы регуляции промоторов.

Репрессор фага λ

Репрессор фага λ представляет собой димер из двух идентичных субъединиц. Молекула димера связывается с участком ДНК длиной 17 п. н. Субъединицы состоят из двух компактных доменов примерно одинаковых размеров, соединенных короткой перемычкой (всего в мономере репрессора 236 а. о.). Перемычка легко расщепляется протеазами. N-концевой домен, полученный после расщепления протеазами (остатки 1—92), способен связываться с ДНК, однако менее прочно, чем интактный репрессор. Кроме того, изолированные N-концевые домены легко диссоциируют до мономеров. Поэтому считается, что в непосредственном контакте с ДНК находится только N-концевой домен, а C-концевой домен стабилизирует связывание с ДНК за счет стабилизации димера.

В N-концевом домене имеется пять участков α -спирали. Первые четыре спирали вместе с соединительными участками образуют компактную глобулу. Пятая спираль (нумерация идет от N-конца) выступает из глобулы и примыкает к пятой спирали другого мономера. За счет этого контакта N-концевые домены удерживаются в димере. Из глобулы выступают также первые 8 а. о., образующих вытянутую «руку».

Рассмотрение пространственной структуры сразу дает представление об общем характере взаимодействия репрессора с оператором. В димере два одинаковых α -спиральных участка 3 образуют выступающие кромки, расстояние между центрами которых составляет 3,4 нм, т. е. равно расстоянию между двумя соседними большими бороздками двойной спирали ДНК. α -Спирали репрессора наклонены так, что могут свободно войти в две соседние большие бо-

роздки оператора, в которых боковые радикалы аминокислотных остатков могут непосредственно контактировать с химическими группами азотистых оснований, выступающими в эту бороздку. Спираль 2 располагается над бороздкой, и ее концевая часть может контактировать с фосфатным остовом молекулы. Гибкая «рука» репрессора (первые 8 аминокислот) обнимает молекулу ДНК (рис. 87). Основную роль в узнавании нуклеотидной последовательности оператора играет спираль, погруженная в большую бороздку ДНК, а остальные контакты неспецифически стабилизируют связь репрессора с оператором. Наиболее убедительно об этом свидетельствуют изящные опыты по направленному изменению специфичности репрессора фага 434, родственного фагу λ . Ген репрессора фага 434 был перестроен так, что аминокислотные остатки, расположенные на наружной поверхности α -спирали 3, входящей в большую бороздку оператора фага 434, были заменены на соответствующие аминокислотные остатки репрессора другого фага P22. Аминокислотные остатки, расположенные на внутренней стороне этой спирали и контактирующие с телом репрессора, были оставлены неизменными. Полученный гибридный репрессор потерял способность связываться с оператором фага 434 и приобрел способность связываться с оператором фага P22.

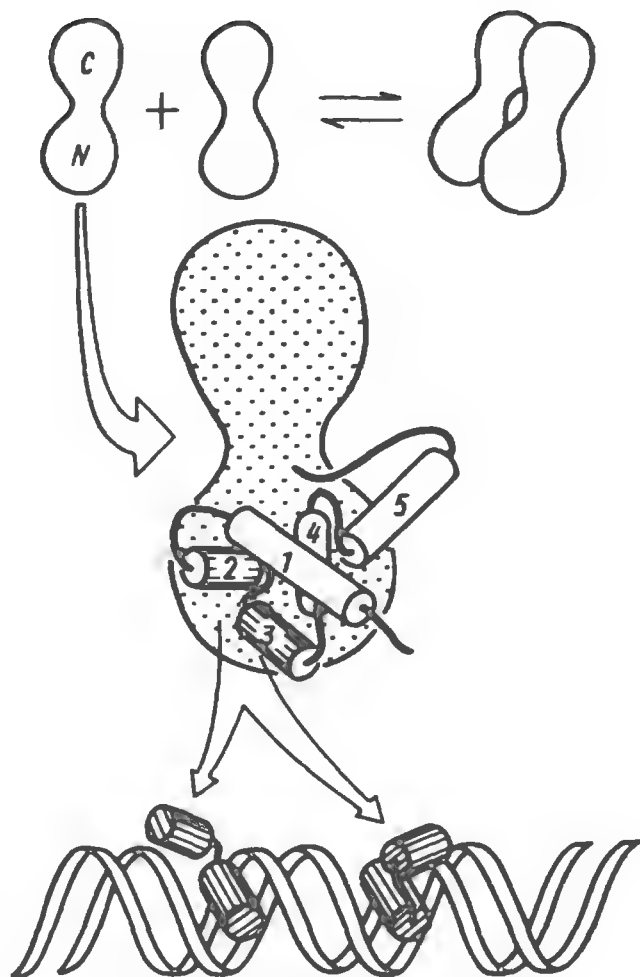


Рис. 87. Репрессор фага λ :
1—5 — номера α -спиралей

Репрессор регулирует активность трех промоторов фага, из которых два, P_{RM} и P_R , располагаются рядом. Транскрипция с промоторов P_{RM} и P_R идет в противоположных направлениях. Между стартовыми точками этих промоторов располагаются три участка связывания репрессора: O_{R1} , O_{R2} и O_{R3} (рис. 88). Участок O_{R3} перекрывается с участком связывания РНК-полимеразы с промотором P_{RM} , поэтому связывание репрессора с O_{R3} мешает связыванию РНК-полимеразы с P_{RM} и тем самым подавляет транскрипцию.

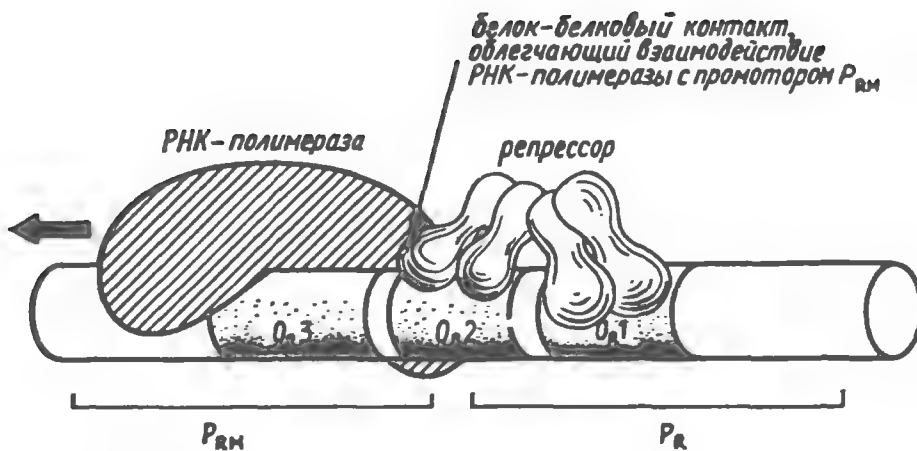


Рис. 88. Активация промотора P_{RM} репрессором фага λ .

Точно так же связывание репрессора с O_{R1} подавляет транскрипцию P_R .

Последствия связывания с O_{R2} более сложные. Этот участок частично перекрывается с участком связывания РНК-полимеразы с промотором P_R , поэтому связывание репрессора с O_{R2} подавляет транскрипцию с P_R . Степень перекрывания O_{R2} с участком связывания РНК-полимеразы с промотором P_{RM} очень мала: в нем имеется только одна фосфатная группа ДНК, с которой контактируют и РНК-полимераза, и репрессор. Поэтому можно думать, что связывание репрессора с O_{R2} не мешает связыванию РНК-полимеразы с P_{RM} . Более того, показано, что репрессор, связываясь с O_{R2} , значительно стимулирует (до десяти раз) транскрипцию с P_{RM} . Предполагается, что активирующее действие репрессора обусловлено тем, что в районе «общего» фосфата между РНК-полимеразой и репрессором возникает белок-белковый контакт, помогающий РНК-полимеразе начать транскрипцию с промотора P_{RM} (рис. 88).

Наиболее убедительным доводом в пользу предположения о специфическом контакте репрессора и РНК-полимеразы служит анализ мутантных форм λ -репрессора, которые сохранили способность связываться с оператором, но потеряли способность стимулировать транскрипцию с P_{RM} . Оказалось, что эти мутации затрагивают биспиральный ДНК-узнающий элемент репрессора. Измененные аминокислотные остатки не участвуют в контактах с ДНК и, располагаясь на поверхности репрессора, могут контактировать с находящимся поблизости участком РНК-полимеразы (рис. 88).

Три операторных участка O_{R1} , O_{R2} , O_{R3} несколько отличаются по нуклеотидной последовательности. Отличается и сродство репрессора к этим участкам. Однако картина распределения молекул репрессора на трех операторных участках зависит не только от относительного сродства репрессора к каждому из этих участков, взятых по отдельности, но и от взаимодействия между димерами репрессора, присоединившимися к соседним участкам. Наибольшим

сродством репрессор обладает к участку O_{R1} . Сродство репрессора к соседнему O_{R2} примерно в десять раз ниже, однако даже при не очень высокой концентрации репрессора участки O_{R1} и O_{R2} оказываются занятыми одновременно за счет того, что связывание репрессора с O_{R1} облегчает связывание другой молекулы репрессора с O_{R1} . Взаимодействие молекул репрессора является парным: контакт с одним соседом исключает взаимодействие с другим соседом, поэтому связывание двух молекул репрессора с O_{R1} и O_{R2} не помогает связыванию с участком O_{R3} , который оказывается занятым лишь при высоких концентрациях репрессора. Если повредить с помощью мутации O_{R1} так, что он потеряет способность связывать репрессор, то взаимодействие репрессоров станет возможным в участках O_{R2} и O_{R3} и они будут заполняться одновременно. Взаимодействие молекул репрессора друг с другом осуществляется через С-концевые домены, так как молекулы, лишенные их, связываются с операторными участками независимо друг от друга.

Следует отметить, что кооперативность взаимодействия репрессоров с операторными участками может проявляться и в том случае, если сконструировать молекулы ДНК, в которых операторные участки удалены друг от друга. Предполагается, что и в этом случае две молекулы репрессора непосредственно контактируют между собой своими С-концевыми доменами, в результате чего ДНК образует петлю. Такие петли легко образуются, когда операторы разделены целым числом витков спирали. В этом случае молекулы репрессора располагаются по одной стороне ДНК, так что для образования петли ДНК нужно только изогнуть (рис. 89). Если расстояние между операторами составляет 2,5 или 3,5 витка спирали, образование петли затруднено, так как молекулы репрессора оказываются на разных сторонах ДНК. Поэтому для образования петли ДНК нужно не только изогнуть, но и закрутить или раскрутить на 180° . Энергия, затрачиваемая на это, весьма ощутима.

Кооперативность взаимодействия репрессора с операторами приводит к тому, что зависимость степени репрессии промотора P_R от концентрации репрессора имеет характерный «сигмоидный» вид (рис. 90). При высоких концентрациях репрессора заняты оба участка, O_{R1} и O_{R2} , что обеспечивает сильную (1000-кратную) репрессию. При незначительном уменьшении концентрации репрессора занятость участков O_{R1} и O_{R2} практически не уменьшается и степень репрессии не меняется. При дальнейшем уменьшении концентрации репрессора степень репрессии резко падает. Таким образом, промотор либо очень сильно репрессирован, либо почти полностью дерепрессирован: для перехода от 1000-кратной до 50 %-ной репрессии требуется уменьшить концентрацию репрессора всего в пять раз.

В отсутствие кооперативности зависимость степени репрессии от концентрации репрессора более плавная. Соответственно для тех промоторов, которые регулируются по принципу резкого переключения, используется кооперативный способ взаимодействия

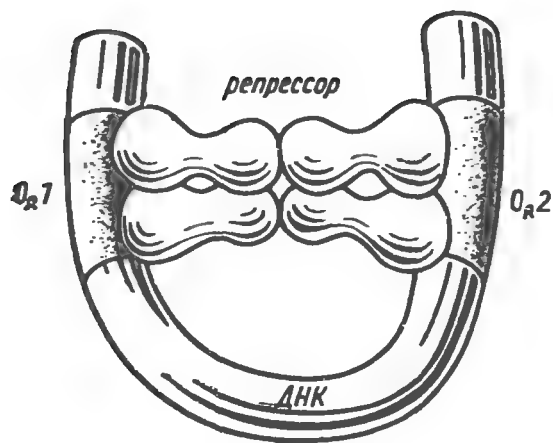


Рис. 89. Образование петли ДНК при кооперативном взаимодействии репрессора с операторными участками, разделенными целым числом витков спирали ДНК

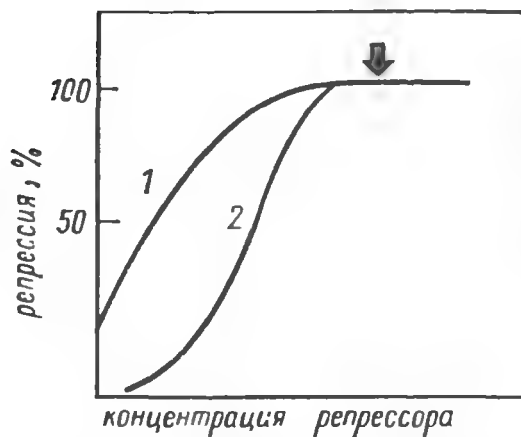


Рис. 90. Зависимость активности промотора от концентрации репрессора при некооперативном (1) и кооперативном (2) взаимодействии репрессора с оператором

репрессора с оператором, состоящим минимум из двух участков связывания. В случае фага λ такой способ переключения предохраняет клетку от случайного запуска развития фага, но легко позволяет включить такое развитие, когда это требуется. В тех случаях, когда необходим плавный переход через промежуточные степени активности, используется некооперативное связывание с одним операторным участком (рис. 90).

Белок-активатор катаболических оперонов

Белок-активатор катаболических оперонов (БАК) в комплексе с циклическим сАМР активирует транскрипцию большого числа оперонов, отвечающих за расщепление различных соединений, преимущественно сахаров, используемых бактериальной клеткой в качестве источников энергии и углерода. Концентрация сАМР в клетках повышается при росте на плохо усваиваемых источниках, например ацетате или глицерине, и снижается при росте на легко усваиваемых, например глюкозе. Поэтому система регуляции с помощью БАК-сАМР позволяет клетке включать опероны катаболизма лишь по мере истощения более легко усваиваемых пищевых веществ.

БАК состоит из двух идентичных субъединиц. Каждая субъединица образует два домена, но в отличие от репрессора фага λ за связывание с ДНК отвечает С-концевой домен. N-концевой домен БАК связывается с сАМР и обеспечивает межсубъединичные контакты в димере. Как и у репрессора фага λ и многих других регуляторов транскрипции, связывание БАК с ДНК осуществляется за счет пары биспиральных элементов, погруженных в большие борозд-

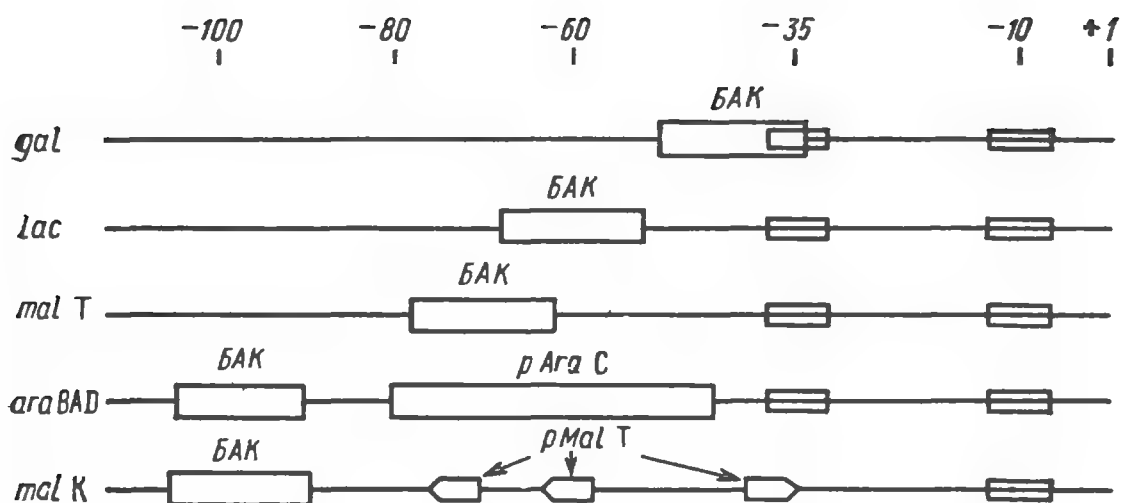


Рис. 91. Расположение участков связывания БАК в промоторах некоторых катаболитных оперонов

ки В-формы ДНК. Особенностью взаимодействия БАК с ДНК является то, что он изгибает ДНК в месте своего присоединения к ней.

Участки связывания БАК в разных промоторах не идентичны по нуклеотидной последовательности, но проявляют значительное сходство друг с другом. Эти участки в разных промоторах располагаются по-разному относительно стартовых точек транскриптов, которые стимулируются БАК. В промоторе галактозного оперона участок связывания БАК непосредственно прилегает к участку связывания РНК-полимеразы. Рядом располагается второй участок связывания БАК; с ним БАК связывается только в том случае, если на промоторе уже находится РНК-полимераза и БАК на первом участке. В промоторе лактозного оперона имеется только один участок связывания БАК, ориентированный относительно участка связывания РНК-полимеразы точно так же, как второй участок связывания БАК в галактозном промоторе. В промоторе гена *malT*, кодирующего белок-активатор мальтозных оперонов, и в промоторах мальтозных и арабинозного оперонов участок связывания отстоит от стартовой точки транскрипции еще дальше (рис. 91). В промоторах мальтозных и арабинозного оперонов, между участками связывания БАК и участками связывания РНК-полимеразы, располагаются участки связывания белков-активаторов этих оперонов, кодируемых генами *malT* и *araC* соответственно.

Механизм действия БАК не вполне понятен. По аналогии с репрессором фаза 1 можно предположить, что существенную роль играют контакты белков-регуляторов между собой и с РНК-полимеразой. Скорее всего с РНК-полимеразой непосредственно взаимодействует лишь ближайший к ней белок. В пользу этого говорит, например, то, что повышение концентрации белков *MalT* и *AraC* снижает зависимость транскрипции соответствующих оперонов от БАК.

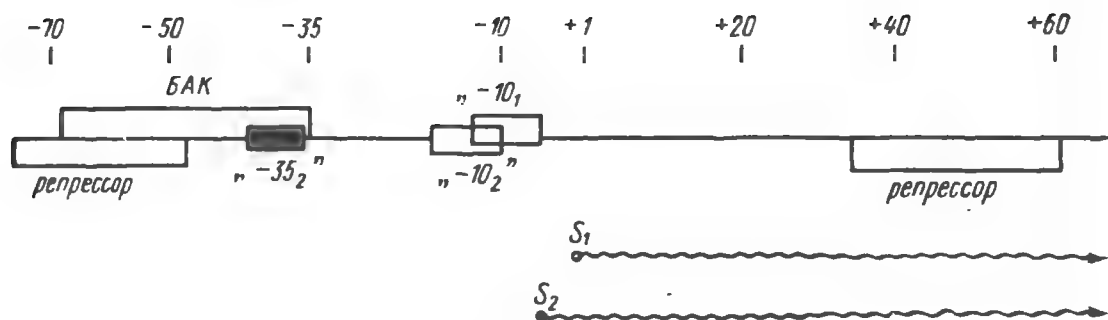


Рис. 92. Промотор галактозного оперона

БАК может выступать также и в роли репрессора транскрипции. Например, в галактозном опероне кроме стимулируемого БАК промотора P_1 имеется репрессируемый БАК промотор P_2 . Эти два промотора перекрываются друг с другом, так что присоединение одной молекулы РНК-полимеразы к промотору P_2 препятствует присоединению другой молекулы РНК-полимеразы к P_2 (рис. 92). Присоединение БАК к ДНК мешает связыванию РНК-полимеразы с P_2 и не мешает связыванию с P_1 . Поэтому БАК оказывает не только прямое, но и опосредованное активирующее действие на промотор P_1 . Блокирование промотора P_2 приводит к усилению транскрипции с P_1 , так как обеспечивает беспрепятственное связывание РНК-полимеразы с P_1 .

Репрессоры катаболитных оперонов

Помимо общей регуляции с помощью БАК-сАМР существует индивидуальная регуляция катаболитных оперонов. Классическим примером является негативная регуляция лактозного оперона. В отличие от ранее рассмотренных димерных белков-регуляторов репрессор лактозного оперона представляет собой тетрамер и содержит два идентичных центра связывания ДНК. Пространственная структура этих центров формируется N-концевыми участками полипептидных цепей, которые, судя по их аминокислотной последовательности, способны образовывать биспиральные элементы, аналогичные биспиральным ДНК-узнающим элементам репрессора фага λ и БАК. С-концевые домены субъединиц лактозного репрессора формируют два центра связывания индуктора лактозного оперона.

Оператор лактозного оперона располагается сразу за стартовой точкой транскрипции. Долгое время считалось, что присоединение лактозного репрессора к промотору стерически мешает присоединению РНК-полимеразы. Однако недавно получены данные, свидетельствующие о том, что репрессор и РНК-полимеразы могут расположиться на промоторе рядом друг с другом. Поэтому приходится думать о более изощренных механизмах репрессии, включающих специфические контакты репрессора с РНК-полимеразой. В лактозном опероне имеется два псевдооператора, сходных по нуклеотидной последовательности с оператором, но обладающих

меньшим сродством к репрессору. Один из этих псевдооператоров располагается внутри гена, кодирующего лактозный репрессор, на расстоянии 92 п. н. от главного оператора, другой — на расстоянии 401 п. н. внутри гена β -галактозидазы. Наличие псевдооператоров несколько увеличивает сродство репрессора к оператору. Стабилизация комплекса репрессора с оператором может происходить за счет того, что одна молекула репрессора, имеющая два центра связывания ДНК, может одновременно присоединить к себе и оператор, и псевдооператор. Разделяющая их ДНК при этом образует петлю.

Два оператора имеется в галактозном опероне. Один из них располагается в районе -60 п. н. промотора, другой — в районе $+55$ (рис. 92). Показано, что связывание репрессора с операторами не мешает связыванию БАК и РНК-полимеразы с промотором. Поскольку для эффективной репрессии нужны оба оператора, предполагается, что молекулы репрессора, расположенные на операторах, взаимодействуют друг с другом, образуя петлю ДНК. Такая конформация каким-то образом мешает инициации транскрипции.

Образование петель постулировано и при репрессии арабинозного оперона *araBAD*. Репрессором этого оперона является белок, кодируемый геном *araC*. В отсутствие арабинозы АгаС-белок, являющийся димером, репрессирует *araBAD*-оперон, а в присутствии арабинозы превращается в активатор, который активирует этот оперон. Кроме того, АгаС-белок как в присутствии, так и в отсутствие арабинозы умеренно репрессирует транскрипцию своего собственного гена, в результате чего концентрация АгаС-белка поддерживается на постоянном уровне.

Ген *araC* располагается перед началом *araBAD*-оперона, но транскрибируется в противоположном направлении (рис. 93).

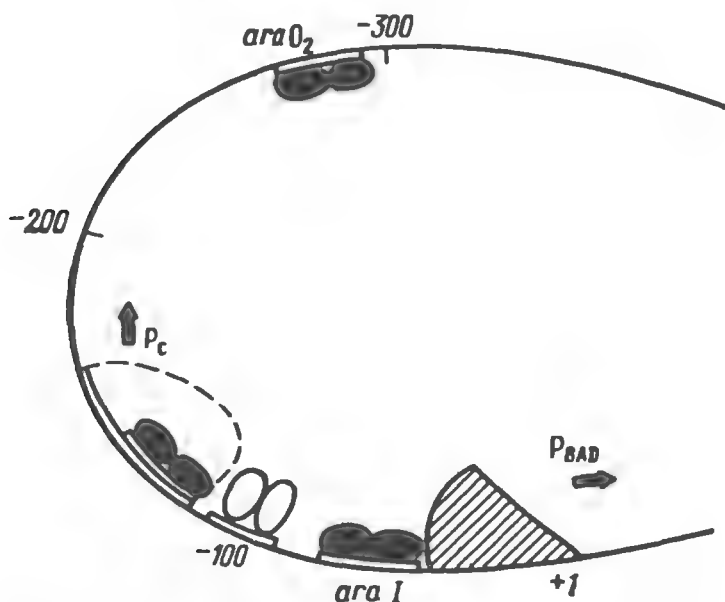


Рис. 93. Промоторная зона арабинозного оперона

Стартовые точки транскрипции промоторов P_C и P_{BAD} находятся на расстоянии 147 п. н. Репрессируя P_C , белок AgaC связывается оператором $araO_1$, перекрывающимся с участком связывания РНК-полимеразы. Находясь в состоянии активатора, AgaC -белок связывается с участком $araI$, непосредственно примыкающим к участку связывания РНК-полимеразы с промотором. Можно думать, что активация происходит за счет контакта с РНК-полимеразой.

Репрессия $araBAD$ -оперона осуществляется в результате связывания AgaC -белка с участком $araO_2$, расположенным на расстоянии почти 300 п. н. от стартовой точки транскрипции (рис. 93). Предполагается, что белок AgaC , располагаясь на $araO_2$, взаимодействует с промоторной зоной за счет изгибания ДНК и образования петли, подобной той, которая образуется при кооперативном взаимодействии двух молекул репрессора фага λ , присоединяющихся к двум операторным участкам, разделенных целым числом витков двойной спирали ДНК. Эта гипотеза основана на результатах экспериментов, в которых изменяли расстояние между $araO_2$ и P_{BAD} за счет небольших делеций и вставок. Когда изменение было кратно одному витку спирали ДНК, степень репрессии не изменялась. Если изменение составляло полвитка, полтора и т. д., то репрессия ослаблялась. Таким образом, для репрессорного действия AgaC -белка существенно не расстояние от промотора, а осевой угол между $araO_2$ и промотором. С каким белком, находящимся на промоторе, контактирует AgaC -белок, находясь на $araO_2$, пока не установлено. Это может быть другая молекула AgaC -белка, расположенная на $araI$, или непосредственно РНК-полимераза.

σ -Субъединицы РНК-полимеразы

σ -Субъединицы РНК-полимеразы можно рассматривать как белки-регуляторы, включающие определенные группы генов. От других белков-активаторов σ -субъединицы отличаются лишь в деталях механизма действия: если «обычные» белки-активаторы сначала присоединяются к промотору, а затем взаимодействуют с РНК-полимеразой, то σ -субъединица сначала присоединяется к РНК-полимеразе, а затем присоединяется к промотору как часть ДНК-связывающего центра фермента.

Впервые включение генов за счет появления в клетке новых σ -субъединиц было продемонстрировано при таком запрограммированном во времени необратимом процессе, как развитие бактериофагов (см. гл. XIII). Другим запрограммированным во времени необратимым процессом, при котором происходит последовательное включение и выключение больших групп генов с участием различных σ -субъединиц, является споруляция таких бактерий, как *Vac. subtilis*.

σ -Субъединицы участвуют также в регуляции ряда промоторов, активность которых обратимо реагирует на изменения внешних

условий, например на тепловой шок. Реакция теплового шока характерна для всех изученных организмов, как эу-, так и прокариот. При резком повышении температуры происходит значительное (но кратковременное) ускорение синтеза определенных белков за счет усиления транскрипции их генов. Промоторы соответствующих генов у *E. coli* отличаются по нуклеотидной последовательности от промоторов, используемых РНК-полимеразой, содержащей главную σ^{70} -субъединицу. Транскрипцию с промоторов, активируемых при тепловом шоке, осуществляет РНК-полимераза, содержащая особую σ^{32} -субъединицу с $M_r=32$ кД. Молекулы σ^{32} отличаются коротким временем жизни в клетке, т. е. расщепляются протеазами через несколько минут после своего образования. Количество σ^{32} в клетке при тепловом шоке возрастает за счет неизвестного еще механизма, что является причиной ускорения транскрипции генов белков теплового шока.

Особая σ -субъединица участвует в транскрипции ряда генов, ответственных за метаболизм азота. К ним относятся ген, кодирующий глутаминсинтетазу, и гены, контролирующие фиксацию атмосферного азота. Промоторы этих генов не содержат обычных для других промоторов последовательностей «—10» и «—35». Вместо них имеются участки гомологии, центры которых расположены в положениях «—11» и «—21». Поэтому неудивительно, что эти промоторы не используются РНК-полимеразой, содержащей главную сигма-субъединицу, σ^{70} . Транскрипцию этих промоторов обеспечивает одна из минорных σ -субъединиц, σ^{60} , кодируемая геном *groN*. Однако для функционирования промотора гена глутаминсинтетазы белка σ^{60} недостаточно. Необходим еще ДНК-связывающийся белок, называемый NR_1 . Перед промотором имеется пять участков его связывания; наибольшее сродство NR_1 проявляет к двум отдаленным участкам. Эти последовательности необходимы для активации промотора при низких концентрациях NR_1 и не обязательны при высоких. Если эти последовательности отодвинуть на тысячу пар нуклеотидов от промотора, они продолжают обеспечивать активность промотора. Предполагается, что белок NR_1 взаимодействует с РНК-полимеразой, расположенной на промоторе. Посадка NR_1 на ДНК облегчает это взаимодействие, сопровождаемое, по-видимому, образованием петли ДНК.

Промотор гена глутаминсинтетазы замечателен не только тем, что он регулируется с участием минорной сигма-субъединицы и нуклеотидных последовательностей, удаленных на большие расстояния от старта транскрипции, но и тем, что действие регуляторного белка модулируется не путем связывания лигандов-эффекторов, которыми могли бы быть глутамин или глутаминовая кислота, а путем химической модификации — фосфорилирования и дефосфорилирования NR_1 , — осуществляемой несколькими ферментами, реагирующими на обеспеченность клетки источниками азота.

Во всех до сих пор рассмотренных примерах регуляции транскрипции на взаимодействие РНК-полимеразы с промотором влияли белки. Регуляция синтеза рибосомных РНК дает пример того, что с РНК-полимеразами могут непосредственно реагировать и низкомолекулярные эффекторы.

В хромосоме *E. coli* имеется 7 копий генов рибосомных РНК. Эти гены организованы в 7 оперонов и располагаются в каждом из них по направлению транскрипции в последовательности 16S—23S—5S. Перед каждым опероном располагается по два промотора, P1 и P2, отстоящих друг от друга на расстоянии более 100 п. н. P1 является сильным промотором, P2 — относительно слабым.

Синтез рибосомных РНК строго скоординирован с синтезом рибосомных белков так, что в клетках в заметных количествах не обнаруживается ни свободных рибосомных РНК, ни свободных рибосомных белков. Скорость образования рибосом регулируется: в быстро растущих на богатых питательных средах культурах эта скорость высокая, в медленно растущих на бедных средах — низкая. Механизмы координированной регуляции синтеза компонентов рибосом отличаются большой сложностью и изучены еще недостаточно. Здесь будет рассмотрен только один элемент этой регуляции, основанной на взаимодействии с РНК-полимеразой низкомолекулярного эффектора гуанозинтетрафосфата. Этот нуклеотид синтезируется на рибосомах в условиях аминокислотного голодания клеток. Накопление гуанозинтетрафосфата в голодающих по аминокислотам клеткам приводит к значительному замедлению синтеза рибосомных РНК и мРНК рибосомных белков и может стимулировать транскрипцию оперонов биосинтеза аминокислот.

Добавление гуанозинтетрафосфата к очищенной РНК-полимеразе подавляет транскрипцию оперона рибосомной РНК с промотора P1, но не влияет на транскрипцию с P2. Поэтому подавление транскрипции рибосомной РНК гуанозинтетрафосфатом никогда не бывает полным.

В пользу того, что и в живых клетках гуанозинтетрафосфат действует непосредственно на РНК-полимеразу, свидетельствует получение мутаций, изменяющих ее β -субъединицу и приводящих к нечувствительности синтеза РНК к гуанозинтетрафосфату.

4. РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ В ТЕРМИНАТОРАХ ПРОКАРИОТ

Очевидная роль терминаторов транскрипции состоит в прекращении синтеза РНК в концах оперонов, что обеспечивает независимую регуляцию экспрессии различных участков ДНК. Но терминаторы встречаются и внутри оперонов. Эффективность этих «внутренних» терминаторов может регулироваться, что позволяет клетке изменять

скорость синтеза РНК на участках ДНК, расположенных за терминаторами, не меняя скорости синтеза РНК на участках, расположенных перед терминаторами. Прежде чем переходить к обсуждению функционального смысла и механизмов такой регуляции, в которой участвуют специальные белковые факторы и рибосомы, рассмотрим процесс терминации в отсутствие факторов.

В их отсутствие РНК-полимераза способна терминировать синтез РНК лишь на некоторых терминаторах, нуклеотидная последовательность в районе которых отличается двумя характерными особенностями. В них по ходу транскрипции сначала идет GC-богатый участок, обладающий центральной симметрией, а затем участок из 4—8 расположенных подряд А в значащей нити. Транскрипция заканчивается на конце олигоА последовательности или сразу за ней. Предполагается, что после прохождения РНК-полимеразой GC-богатого участка с центральной симметрией в РНК-продукте возникает шпилька, приводящая к остановке РНК-полимеразы и разрушению части РНК-ДНК гибрида транскрибирующего комплекса. Оставшаяся часть РНК-ДНК гибрида, содержащая концевую олигоU последовательность РНК, легко плавится ввиду относительной нестабильности $gU \cdot dA$ -пар, что приводит к освобождению РНК-продукта (рис. 94). Первым из комплекса освобождается РНК-продукт, а затем РНК-полимераза. В какой момент происходит схлопывание нитей ДНК, пока не известно.

Эффективность терминации зависит от прочности терминаторной шпильки в РНК. Это видно из того, что мутации, приводящие к нарушению комплементарного спаривания какого-либо основания в шпильке, ослабляют терминацию, а мутации, восстанавливающие комплементарность, усиливают терминацию. Усиление прочности РНК-ДНК гибрида в районе олигоU ослабляет терминацию. Такое усиление происходит при уменьшении длины олигоА в матрице (минимальный размер этой последовательности, при котором терминация еще возможна, четыре нуклеотида).

Таким образом, эффективность терминации зависит, по-видимому, от баланса стабильности различных РНК: РНК и РНК : ДНК двойных спиралей. В энергетику этого баланса вносит вклад и РНК-полимераза, так как определенные мутации, затрагивающие РНК-полимеразу, влияют на эффективность терминации. Эти же мутации влияют и на продолжительность пауз транскрипции, что подтверждает представление о том, что подготовительной стадией терминации является пауза.

Терминация в отсутствие факторов, по-видимому, не полностью моделирует этот процесс в том виде, как он протекает в живой клетке. Действительно, на ряде терминаторов очищенная РНК-полимераза завершает синтез молекул РНК, отступя на несколько нуклеотидов от того места, где она завершает его в живой клетке, и с меньшей эффективностью. Недавно обнаружен белковый фактор, названный *tau*(τ), который улучшает эффективность и точность

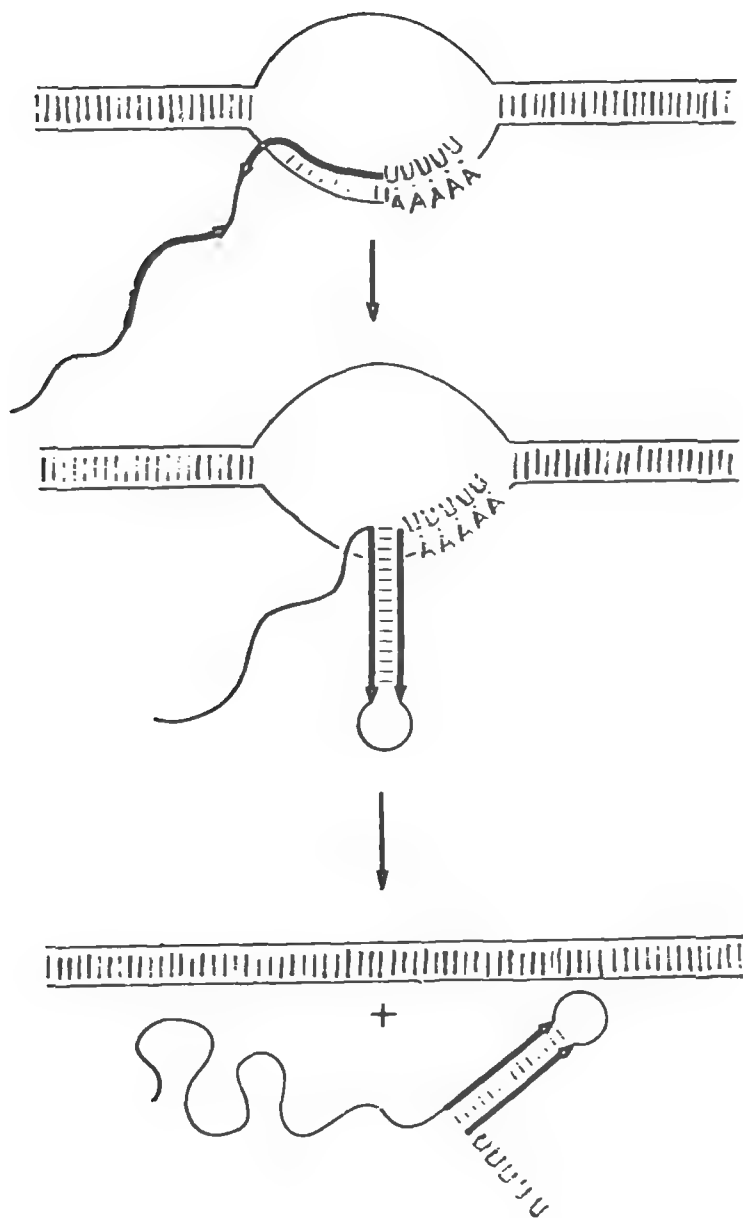


Рис. 94. Терминация транскрипции в результате образования шпильки в РНК

терминации на таких терминаторах. Механизм его действия еще не изучен.

Многие терминаторы узнаются РНК-полимеразой только с помощью фактора терминации, названного ρ . Этот белок с молекулярной массой 46 кД обладает РНК-зависимой нуклеозидтрифосфатазной активностью, обязательной для завершающего действия. НТФазная активность фактора ρ проявляется только в комплексе с одонитевой РНК. Наибольшей НТФазной активностью ρ -фактор обладает в присутствии полицитидиловой кислоты. Фактор ρ агрегирует с образованием гексамера, способного связываться с РНК. В комплексе с РНК гексамер защищает в ней от действия

РНКазы 80—90 нуклеотидов, так что каждый мономер связывает по 12—14 нуклеотидов РНК.

Фактор ρ присоединяется к РНК-продукту до того, как РНК-полимераза достигает терминатора. Присоединение происходит к определенным участкам РНК, в нуклеотидной последовательности которых пока не обнаружено каких-либо характерных особенностей. Ясно лишь, что эти участки не склонны к образованию протяженных двуспиральных структур.

В местах ρ -зависимой терминации РНК-полимераза делает паузы в отсутствие ρ -фактора, поэтому считается, что роль ρ -фактора заключается в вытеснении РНК из транскрипционного комплекса в местах пауз. Рассматриваются две модели. Согласно одной из них, фактор движется по синтезируемой РНК, а в местах пауз догоняет РНК-полимеразу и вытесняет РНК-продукт. Другая модель основана на том, что пирофосфат подавляет НТФазную активность ρ -фактора. Согласно этой модели, ρ -фактор движется за РНК-полимеразой без отставания, но при нормальной скорости элонгации ингибируется пирофосфатом, высвобождающимся при синтезе РНК. Активация ρ -фактора происходит лишь в местах пауз, где синтез цепи РНК временно останавливается, что приводит к прекращению освобождения пирофосфата.

Конкретный механизм вытеснения РНК из транскрипционного комплекса под действием ρ -фактора еще не выяснен. Это вытеснение сопряжено с гидролизом НТФ, при котором, по-видимому, происходит конформационное изменение ρ -фактора. В результате этого изменения РНК вытесняется из комплекса либо непосредственно ρ -фактором, либо за счет воздействия на РНК-полимеразу.

Однако ρ -фактор вызывает терминацию не во всех местах пауз. Например, в случае транскриптов, инициированных на уже известном нам промоторе P_R фага λ , ρ -зависимая терминация происходит лишь в местах пауз, расположенных от промотора на расстояниях, больших 290 н. п., хотя длительные паузы возникают и в более близких к промотору местах. Анализ нуклеотидной последовательности показывает, что начальная часть РНК способна образовывать большое число двунитчатых структур. По-видимому, в данном случае сильно развитая вторичная структура транскрипта мешает связыванию с ним ρ -фактора, без которого терминации в местах пауз не происходит.

Связыванию ρ -фактора с РНК мешают также рибосомы, транслирующие РНК, поэтому на ρ -зависимых терминаторах, встречающихся внутри структурных генов, ρ -фактор не обеспечивает терминации, если мРНК эффективно транслируется. Наличие таких терминаторов внутри генов, по-видимому, не случайно. Когда синтез белка по каким-либо причинам подавлен, они сигнализируют РНК-полимеразе о том, что мРНК не транслируется и синтез ее бессмыслен.

Регулируемые терминаторы бактерий называют *аттенюаторами* (ослабителями). Впервые обнаружен и лучше других изучен аттенюатор триптофанового оперона *E. coli*. Этот оперон состоит из пяти генов, кодирующих ферменты биосинтеза триптофана. Регуляцию осуществляют две системы, чувствующие потребность клетки в триптофане. Первая система влияет на эффективность инициации на промоторе оперона. Репрессор триптофанового оперона в комплексе с триптофаном присоединяется к оператору, расположенному перед стартовой точкой транскрипции в районе «—10», и стерически препятствует РНК-полимеразе присоединиться к промотору. Таким образом, при избытке триптофана оперон репрессируется. В отсутствие триптофана репрессор теряет способность связываться с оператором, в результате чего оперон индуцируется. Эту систему дополняет регуляция в аттенюаторе, расположенном на расстоянии 180 п. н. от стартовой точки транскрипции внутри «лидерной» последовательности, предшествующей иницирующему кодону первого структурного гена. В условиях избытка триптофана лишь одна из десяти молекул РНК-полимеразы, начавших синтез РНК на триптофановом промоторе, преодолевает этот терминатор и переходит в область структурных генов. При уменьшении количества триптофана доля молекул РНК-полимеразы, преодолевающих аттенюатор, возрастает.

Перед аттенюатором располагается несколько областей с центральной симметрией. Образующаяся на них лидерная РНК содержит четыре участка, способных образовывать шпильки в различных сочетаниях (рис. 95). Могут образоваться шпильки 1 : 2 и 3 : 4; шпилька 2 : 3 при этом образоваться не может. Если образуется шпилька 2 : 3, то не могут образовываться шпильки 1 : 2 и 3 : 4. Шпилька 3 : 4 представляет собой терминаторную шпильку, характерную для обычных р-независимых терминаторов, так как после нее в РНК следует несколько U подряд. Поэтому при образовании этой шпильки в аттенюаторе происходит терминация. Эффективность ее составляет около 80—90 %. Если же шпильки 3 : 4 не образуется (за счет образования шпильки 2 : 3), то терминации не происходит.

Образование альтернативных шпилек в лидерной области, от которых зависит терминация в аттенюаторе, определяется тем, как транслируется лидерная РНК (рис. 96). Эта РНК кодирует так называемый лидерный пептид, содержащий два триптофановых остатка подряд. Если этот пептид не транслируется, то вторичная структура РНК формируется по мере ее роста и выхода из транскрипционного комплекса: сначала формируется шпилька 1 : 2, что делает невозможным образование шпильки 2 : 3, а затем шпилька 3 : 4, что приводит к терминации синтеза РНК. Если в клетке имеется недостаток триптофана, то рибосома, транслирующая лидер-

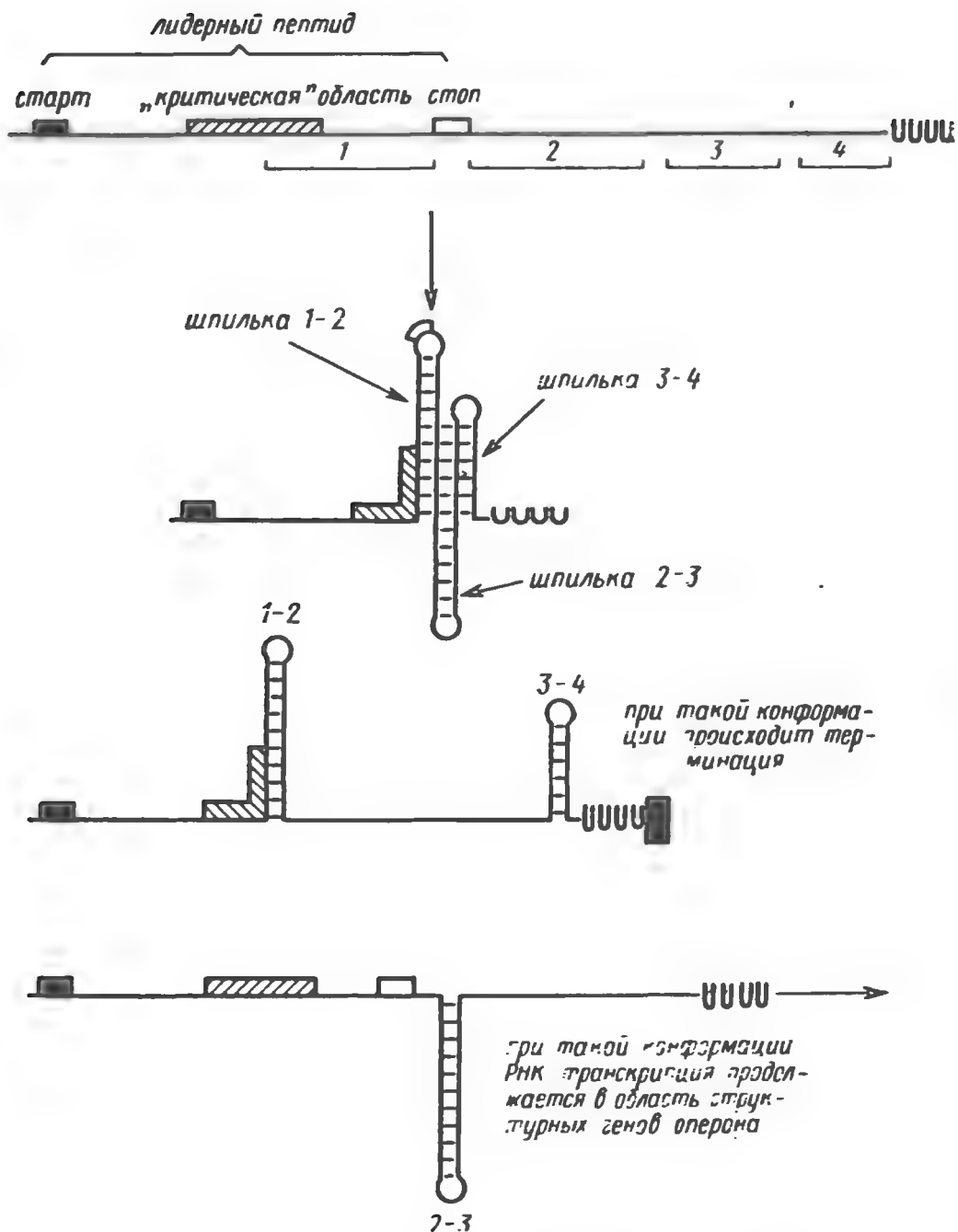


Рис. 95. Атенюатор триптофанового оперона

ный пептид, задерживается на его триптофановых кодонах. Эти кодоны расположены так, что, находясь на них, рибосома прикрывает последовательность 1 и мешает образованию шпильки 1 : 2. В результате становится возможным образование «антитерминаторной» шпильки 2 : 3. Терминации не происходит, и РНК-полимераза переходит в область структурных генов.

Сходная система регуляции в аттенюаторе используется в оперонах, отвечающих за синтез других аминокислот у *E. coli*. В лидерных РНК этих оперонов закодированы специфические пептиды, включающие по несколько остатков той аминокислоты, биосинтез

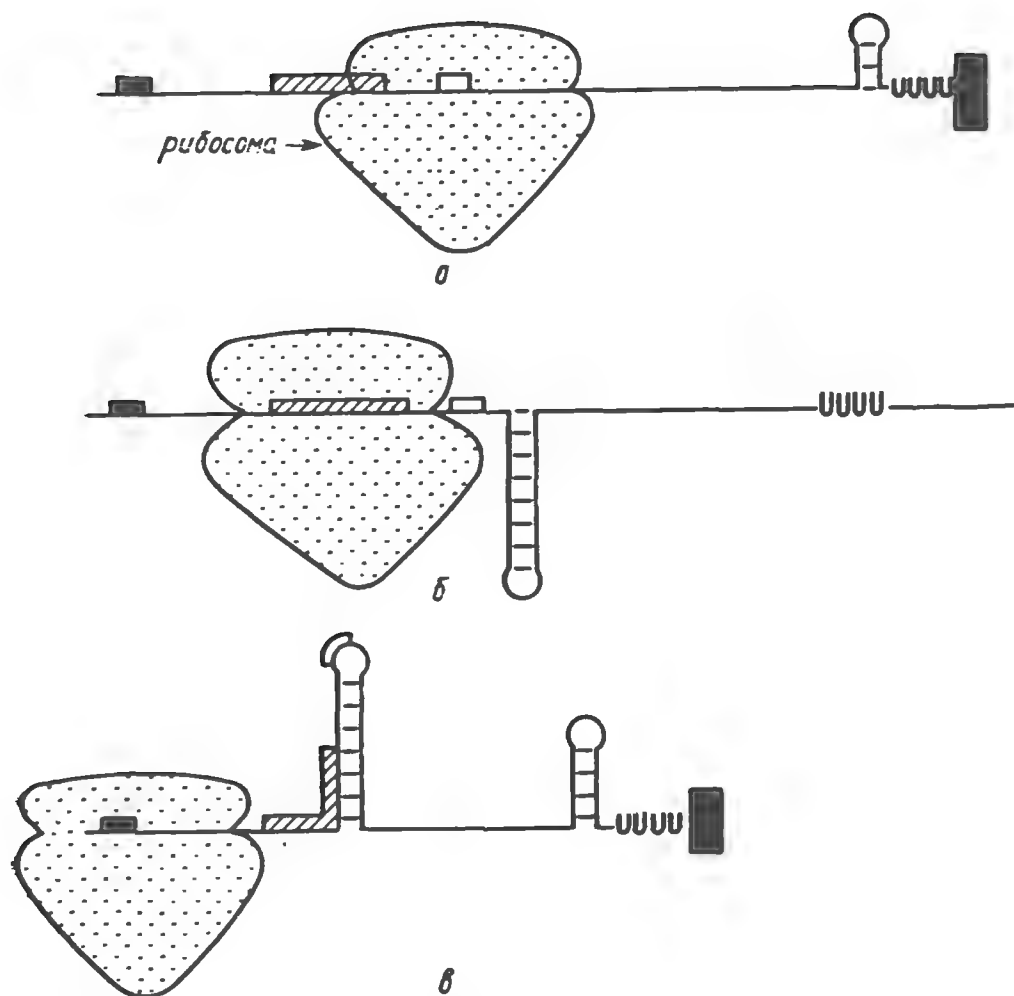


Рис. 96. Влияние рибосомы, транслирующей лидерный пептид, на прохождение РНК-полимеразой аттенюатора:

a — всех аминокислот хватает; *б* — не хватает аминокислоты, за синтез которой отвечает оперон; *в* — не хватает другой аминокислоты или нарушена трансляция

которой определяется опероном. Если в лидерном пептиде триптофанового оперона содержится два tandemно расположенных остатка триптофана, то в лидерном пептиде фенилаланинового оперона — последовательно семь остатков фенилаланина, в которые вклинивается один остаток аланина и один остаток треонина, в лидерном пептиде гистидинового оперона — подряд семь остатков гистидина, в лидерном пептиде лейцинового оперона — четыре остатка лейцина подряд и т. д. При отсутствии соответствующей аминокислоты рибосома задерживается на ее кодонах и обеспечивает образование «антитерминаторной» шпильки.

Аттенюаторы могут регулироваться и в зависимости от уровня нуклеозидтрифосфатов. Например, в *pur*-опероне в лидерной РНК возможно образование двух шпилек, одна из которых является терминаторной. Первая от промотора шпилька вызывает при низкой концентрации УТР паузу транскрипции; в результате рибосома, синтезирующая лидерный пептид, догоняет РНК-полимеразу и

мешает образованию терминаторной шпильки. При высокой концентрации УТР пауза не возникает и первой успевает образоваться антитерминаторная шпилька.

Аттенюатор может быть настроен на количество рибосом в клетке. Например, в случае оперона β -лактамазы *E. coli* ρ -независимый терминатор, расположенный в конце лидерной РНК, перекрывается с участком связывания рибосом. В результате при высокой концентрации рибосом вероятность нарушения структуры терминаторной шпильки больше, чем при низкой. Поскольку концентрация рибосом в быстро делящихся клетках выше, чем в медленно делящихся, синтез β -лактамазы оказывается скоординированным со скоростью деления клеток.

Аттенюатор может регулироваться и без участия рибосом (или, говоря осторожнее, без трансляции лидерной РНК). Так, лидерная РНК триптофанового оперона *Bac. subtilis* не кодирует лидерного пептида. Тем не менее эффективность терминации в аттенюаторе зависит от концентрации триптофана. Предполагается, что баланс в образовании терминаторной и антитерминаторной шпильки в лидерной РНК определяется специальным белком, присоединяющимся к лидерной РНК и блокирующим образование антитерминаторной шпильки.

Своеобразная форма аттенюаторной регуляции обнаружена в *crp*-опероне, кодирующем БАК. Этот оперон авторегулируется, т. е. его транскрипция подавляется БАК в комплексе с сАМР. В районе промотора этого оперона имеется участок связывания БАК. Оказалось, что, связываясь с этим участком, БАК активирует транскрипцию с другого, направленного в противоположную сторону промотора, с которого образуется РНК, комплементарная первым 11 нуклеотидам *crp*-РНК. В результате «антилидерная» РНК образует двунитевую структуру с *crp*-РНК, сходную с терминаторной шпилькой, что приводит к терминации едва начатой с рРНК.

Антитерминация у фага λ и в генах рибосомных РНК

Включение определенных генов фага λ происходит за счет исключения ρ -зависимых и ρ -независимых терминаторов, расположенных перед ними, под действием белка, кодируемого геном N. Кроме фогового N-белка в антитерминации участвуют по крайней мере три белка, кодируемых клеткой-хозяином, из которых более подробно изучен белок, кодируемый геном *nusA*, который способен взаимодействовать с минимальной РНК-полимеразой, занимая на ней место σ -субъединицы. Хотя белок NusA является компонентом системы антитерминации, сам по себе он способен удлинять паузы в некоторых участках ДНК и даже вызывать терминацию.

Белки, участвующие в антитерминации, присоединяются к транскрибирующему комплексу до того, как он достигает терминаторов.

Присоединение белков-антитерминаторов происходит в то время, когда РНК-полимераза проходит специальный сигнал антитерминации, так называемый участок *nut*. Этот участок располагается между промотором и терминатором в нетранслируемой части транскрипта. Предполагается, что в тот момент, когда РНК-полимераза достигает участка *nut*, к ней присоединяется белок Nus A. После этого с комплексом связывается р-фактор. При дальнейшем продвижении РНК-полимеразы в пределах *nut*-участка в комплекс включаются белок N и другие белки, необходимые для антитерминации. В такой «регулосоме» РНК-полимераза приобретает особую конформацию, которая позволяет ей проходить через любые сигналы терминации, расположенные на значительном расстоянии от участка *nut*.

Участки, сходные по последовательности с *nut*-участками фага λ , обнаружены и в некоторых оперонах хромосомы *E. coli*, в частности в оперонах рибосомных РНК. Внутри этих оперонов имеются р-зависимые терминаторы, на которых, однако, в норме терминации не происходит. С низкой эффективностью идет в этих оперонах терминация и на искусственно введенных р-зависимых терминаторах. Терминаторы внутри генов рибосомной РНК, по-видимому, не нужны для его функционирования. Они просто там есть, потому что некоторые последовательности, необходимые для функционирования рРНК, проявляют терминирующие свойства. Когда такие последовательности встречаются в составе мРНК, с терминацией борется рибосома, осуществляющая трансляцию. Поэтому в отсутствие трансляции внутри многих генов происходит терминация транскрипции. В случае рибосомных РНК, которые никогда не транслируются, для борьбы с терминацией в ненужных местах предусмотрена система антитерминации, формирующаяся в начале оперона. Какие белки принимают участие в этой системе, еще не ясно.

ЛИТЕРАТУРА

Никифоров В. Г. РНК-полимераза бактерий: сравнительные исследования. // Усп. микроб. 1987. Т. 21. С. 105—151.

Irwin N., Ptashne M. Mutants of the catabolite activator protein of *Escherichia coli* that are specifically deficient in the gene-activation function // Proc. Natl. Acad. Sci. USA. 1987. V. 81. P. 8315.

Lichenstein H. S., Hamilton E. P., Lee N. Repression and catabolite gene activation in the *araBAD* operon // J. Bacteriol. 1987. V. 169. P. 811.

McClure W. R. Mechanism and control of transcription initiation in prokaryotes // Ann. Rev. Biochem. 1985. V. 54. P. 171.

Platt T. Transcription termination and the regulation of gene expression // Ann. Rev. Biochem. 1986. V. 55. P. 339.

Ptashne M. Gene regulation by proteins acting nearby and at a distance // Nature. 1986. V. 322. P. 697.

Reitzer L. J., Magasanik B. Transcription of *glnA* in *E. coli* is stimulated by activator bound to sites far from the promoter // Cell. 1986. V. 45. P. 785.

Reznikoff W. S., Siegele D. A., Cowing D. W., Gross C. A. The regulation of transcription initiation in bacteria // Ann. Rev. Genet. 1985. V. 19. P. 355.

Roberts J. W. Phage Lambda and the regulation of transcription termination // Cell. 1988. V. 52. P. 5.

Глава VIII

ПРОЦЕССИНГ ПЕРВИЧНЫХ ТРАНСКРИПТОВ

Молекулы предшественников зрелых клеточных РНК подвергаются расщеплению и химической модификации. Совокупность биохимических реакций, в результате которых уменьшается молекулярная масса РНК-предшественника и осуществляются разные способы химической модификации с образованием зрелых молекул РНК, называют *процессингом*. Процессинг наблюдается и в прокариотических клетках, но особенно сложны превращения предшественников клеточных РНК в ядрах эукариот. Хромосомы эукариотической клетки, в которых осуществляется транскрипция, локализованы в ядре и отделены двойной ядерной мембраной от цитоплазмы, где протекает трансляция. В ядре синтезируются предшественники всех типов цитоплазматических РНК. Зрелые молекулы РНК транспортируются в цитоплазму. Механизм транспорта РНК из ядра в цитоплазму исследован недостаточно. Полагают, что процессинг РНК с образованием зрелых молекул продолжается и в ходе их транспорта в составе рибонуклеопroteinных частиц через поры ядерных мембран. В клетках эукариот только незначительная часть, около 10%, транскрибируемых в ядре последовательностей ДНК выявляется в составе цитоплазматических мРНК. Основная часть новообразованной РНК распадается в ядре и не обнаруживается в цитоплазме.

Исследование закономерностей процессинга необходимо для выяснения механизмов регуляции экспрессии генов. Рассмотрение этапов процессинга и его вариантов у разных организмов затрагивает также ряд других принципиальных проблем. Оказалось, что молекула РНК и в отсутствие белка может выступать как аутокатализатор, осуществляя благодаря конформационной гибкости молекулы сложную и точную собственную перестройку с образованием новых ковалентных связей. Таким образом, полирибонуклеотиды могут функционировать подобно ферментам. Открытие возможности аутокаталитических превращений полирибонуклеотидов показало, что исследование процессинга РНК имеет прямое отношение к вопросу о пребиотических стадиях эволюции макромолекул.

1. ПРОЦЕССИНГ У ПРОКАРИОТ

В бактериальных клетках большей частью образуются и транслируются полицистронные мРНК, однако некоторые типы мРНК моноцистронные. Процессинг предшественников мРНК не характерен для бактериальных клеток, хотя случаи процессинга обнаружены, например, при транскрипции оперона, кодирующего рибосомальные белки (L1, L7.12) и субъединицы РНК-полимеразы (β и

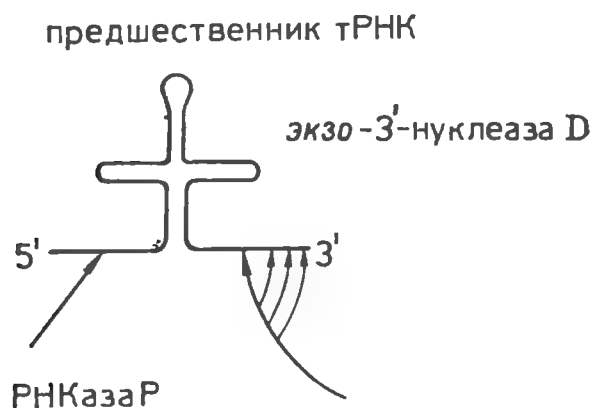
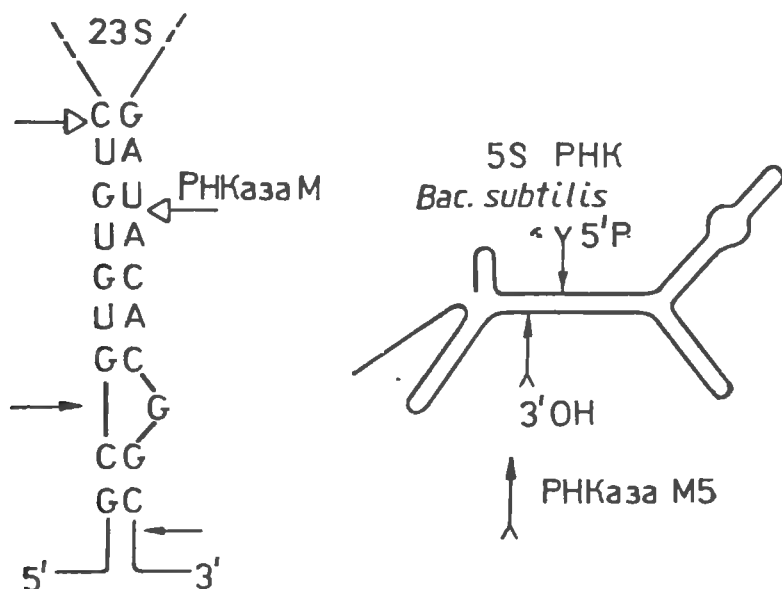
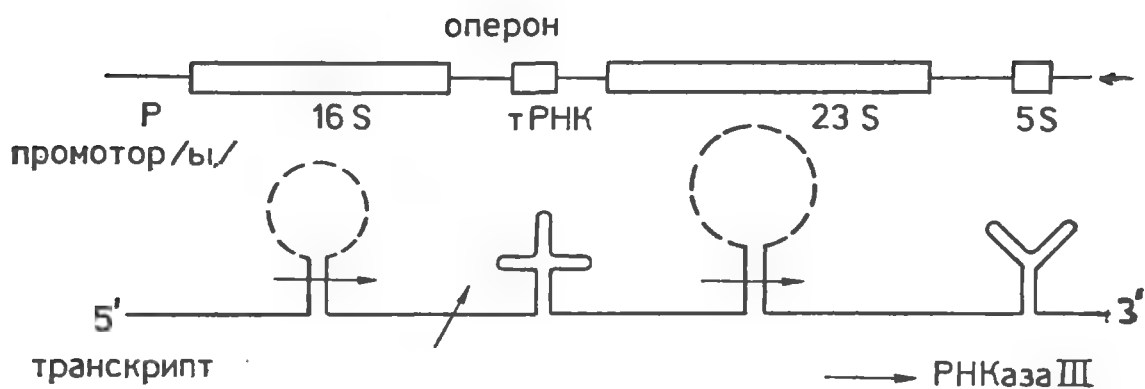


Рис. 97. Процессинг рРНК и тРНК в клетках бактерий

β'). Гены расположены в следующем порядке: $L1 \rightarrow L7/12 \rightarrow \beta \rightarrow \beta'$. Транскрипция идет в направлении от $L1$ к β' . Созревание полицистронной РНК осуществляется за счет эндонуклеазного расщепления РНК, в результате которого разъединяются мРНК для рибосомальных белков и субъединиц РНК-полимеразы. Однако этот процесс,

по-видимому, не является обязательным для обеспечения трансляции проксимальных L1, L7/12 или дистальных (β , β') генов.

Еще недавно считали, что полиаденилирование 3'-конца мРНК происходит только в клетках эукариот (см. раздел 4 этой главы). Теперь показано, что существенная часть мРНК бактерий также подвергается посттранскрипционному полиаденилированию.

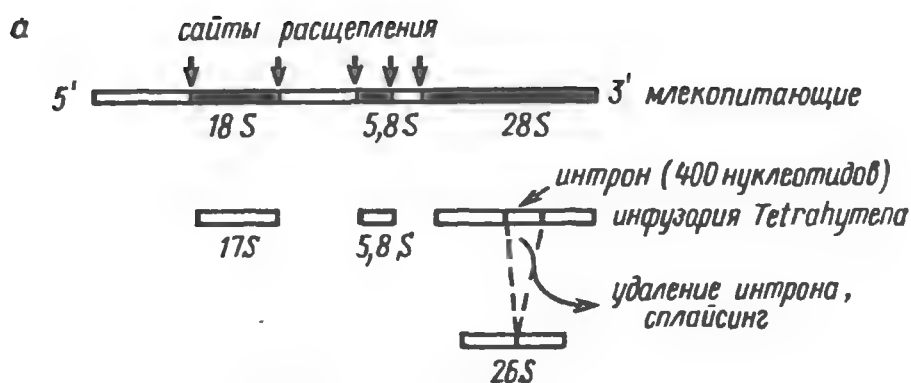
Процессинг необходим в бактериальных клетках при образовании зрелых молекул рРНК и тРНК. Транскрипт-предшественник образуется на оперонах, включающих гены 16S, 23S, 5S и тРНК (рис. 97). В хромосоме *E. coli* находятся 7 таких оперонов, различающихся друг от друга типами и числом генов для тРНК. Транскрипт, включающий около 5600 нуклеотидов, по-видимому, по мере своего образования подвергается действию эндонуклеаз и экзонуклеаз. Специфичность их действия в значительной степени обусловлена особенностями вторичной структуры образующегося транскрипта, схематически представленной на рис. 97.

РНКаза III (эндонуклеаза) атакует двунитевые участки (черешки) петель, включающие последовательности будущих зрелых 16S (1542 нуклеотида) и 23S (1904 нуклеотида) рРНК. В результате осуществляется одновременное созревание 5'- и 3'-концов рРНК. Вслед за РНКазой III, вероятно после взаимодействия транскрипта с рибосомными белками, в процессинге примут участие и другие эндонуклеазы (М-эндонуклеаза, от англ. maturation — созревание). Действие всех эндонуклеаз сопровождается образованием 5'-Р- и 3'-ОН-концов. В структурах, расщепляемых РНКазой III, не обнаружено сходства на уровне нуклеотидной последовательности, за исключением общего требования — двунитевой структуры, которая окружает или включает сайт расщепления. Подобным образом действует и РНКаза, участвующая в процессинге 5S РНК (рис. 97).

Функциональный смысл такого способа организации оперонов рРНК и процессинга транскрипта не совсем ясен. Можно предполагать, что многоступенчатость процессинга позволяет на каждой его стадии регулировать скорость образования зрелых РНК. Кроме того, локализация в одном первичном транскрипте будущих 16S и 23S рРНК, входящих в состав малой и большой субъединиц рибосом, автоматически обеспечивает их образование в равных молярных соотношениях.

2. ПРОЦЕССИНГ рРНК У ЭУКАРИОТ

Гены рРНК у эукариот представлены тандемно повторяющимися копиями (100—400), служащими матрицами для образования транскриптов, подвергающихся процессингу (рис. 98). Транскрибируемые последовательности разделены *спейсерами*, также играющими большую роль в транскрипции рРНК и ее регуляции (см. гл. X,



б Схема реакций при сплайсинге рРНК

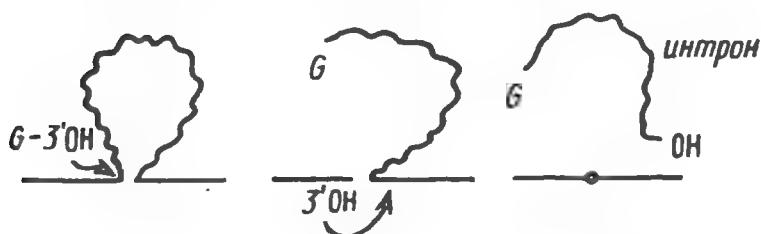
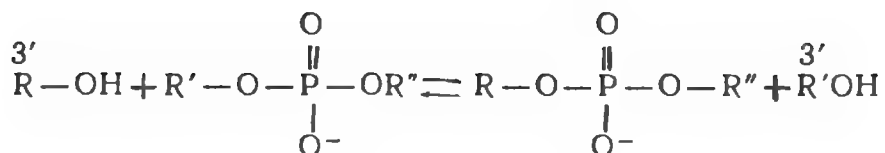


Рис. 98. Процессинг рибосомальных рРНК у эукариот (а) и схема реакций при сплайсинге рРНК (б)

раздел 3). В результате процессинга, в ходе которого метилируется ряд рибозных остатков, из молекулы предшественника выщепляются три типа рРНК: 18S, входящая в состав малой субъединицы рибосом, а также 28S и 5,8S, локализующиеся в большой субъединице.

У некоторых организмов в составе предшественника 28S РНК находится вставка — *интрон*, который удаляется при процессинге. Сшивание фрагментов молекулы РНК при удалении интрона получило название *сплайсинга* (рис. 98). Сплайсинг фрагментов РНК большой рибосомальной субъединицы обнаружен и изучен у ресничной инфузории *Tetrahymena* при функционировании генов «вегетативного ядра» — макронуклеуса, а также при процессинге митохондриальной рибосомальной РНК у дрожжей. Исследование сплайсинга предшественника рРНК инфузорий *in vitro* выявило удивительный результат. Оказалось, что удаление интрона и сплайсинг можно наблюдать не только в изолированных ядрах, но и при инкубации очищенного предшественника рРНК в отсутствие белка. В последнем случае процесс идет медленнее, но достаточно эффективно, если присутствуют ионы магния и гуанозин (или GMP, GDP, GTP). Таким образом, наблюдается аутокаталитический процесс самосплайсинга (англ. selfsplicing), в результате которого в молекуле РНК разрываются одни межнуклеотидные ковалентные связи и устанавливаются новые. Реакция идет по механизму трансэтерификации:



3'-ОН-группа гуанозина атакует фосфодиэфирную связь между 3'-концом первого экзона и первым нуклеотидом интрона, причем гуанозин присоединяется к 5'-концу интрона (рис. 98). На следующем этапе 3'-ОН-группа на конце первого экзона атакует другой конец интрона. Реакция трансэтерификации сопровождается выщеплением интрона, молекула РНК подверглась *самосплайсингу*. Обе стадии сплайсинга осуществляются путем трансэтерификации, причем в итоге число эфирных связей не меняется. Поэтому сплайсинг может осуществляться без внешнего источника энергии, в отсутствие АТР или GTP.

Вырезание интрона происходит очень точно; это обеспечивается наличием сложной вторичной и третичной структуры РНК. Нуклеотидная последовательность интрона с учетом комплементарных взаимодействий отдельных участков может быть представлена в виде достаточно сложной структуры (рис. 99). Сходную структуру имеет интрон предшественника рРНК митохондрий. Замены отдельных нуклеотидов в составе интрона обнаруживают необходимость отдельных элементов его структуры для самосплайсинга. Например, нарушение комплементарности в районе А препятствует сплайсингу. Оказывается, что для правильного сплайсинга необходимы также комплементарные взаимодействия нуклеотидов (вне плоскости рисунка) в элементах Б и В. Замена нуклеотида в районе Б, нарушившая комплементарность и сплайсинг, может быть компенсирована другой нуклеотидной заменой в районе В, если она восстановит комплементарные взаимодействия. Каталитические свойства определяются особой структурой РНК, создаваемой в результате комплементарных взаимодействий.

Нуклеотидные последовательности этих элементов структуры сохраняются в эволюции. Интроны с подобной структурой и со способностью к «самовырезанию» из предшественника обнаружены в генах трех разных клеточных органелл: в ядрах инфузорий, митохондриях дрожжей и хлоропластах растений. Обнаружение самосплайсинга предшественника рРНК показывает, что молекула РНК может выступать как катализатор или, вернее, как квазикатализатор, поскольку в процессе реакции меняется структура РНК. Подобный случай сплайсинга (самосплайсинга) обнаружен и в прокариотических системах при исследовании экспрессии гена тимидилатсинтазы фага Т4, заражающего клетки *E. coli*. Ген, кодирующий тимидилатсинтазу, содержит интрон, выщепляемый при процессинге мРНК.

Способность РНК осуществлять биокатализ ярко проявляется также при исследовании процессинга предшественников тРНК.

3. ПРОЦЕССИНГ ПРЕДШЕСТВЕННИКОВ тРНК У ПРО- И ЭУКАРИОТ

Выщепление предшественника молекулы тРНК из состава первичного транскрипта у бактерий осуществляется с помощью РНКазы III (см. рис. 97). Предшественник тРНК содержит около 100 нуклеотидов, а зрелые тРНК — 70—90 нуклеотидов. Процессинг 5'-конца предшественников (англ. precursors) тРНК осуществляется с помощью РНКазы Р. По-видимому, РНКазы Р осуществляют процессинг 5'-конца всех тРНК бактериальной клетки.

Процессинг 3'-конца осуществляется с помощью РНКазы D (D — digestion), представляющей собой экзо-3'-нуклеазу (см. рис. 97). Это непротессивный фермент, при работе которого комплекс с субстратом диссоциирует после каждого акта отщепления нуклеотида. Скорость отщепления резко падает, когда нуклеаза достигает тринуклеотидной последовательности CCA_{OH} , необходимой для взаимодействия с аминоацил-тРНК-синтетазой при образовании аминоацил-тРНК. При созревании некоторых видов тРНК 3'-конец CCA_{OH} образуется заново в результате последовательных нуклеотидилтрансферных реакций.

Узнавание тРНК-предшественника РНКазой Р осуществляется при ее взаимодействии со зрелой «клеверной» (или L-образной) структурой молекулы тРНК, причем отщепление с 5'-конца нескольких нуклеотидов, варьирующих в разных тРНК, осуществляется с удивительной точностью. Фермент состоит из двух компонентов: РНК и основного белка с $M_r = 17\,000$. Подобный фермент функционирует и при процессинге тРНК у эукариот. РНК-компонент нуклеазы Р *E. coli* (M1 РНК, от англ. maturation — созревание) содержит 400 нуклеотидов, в клетках дрожжей — 250 нуклеотидов и в клетках человека — около 100 нуклеотидов. Оказалось, что в присутствии высоких концентраций магния молекула РНК без белка способна катализировать созревание предшественника тРНК. Ферментативной активностью обладает РНК, а белок сам по себе неактивен, но усиливает каталитическую активность РНК. В его присутствии также расширяется ассортимент расщепляемых *in vitro* предшественников тРНК. В РНК-компоненте РНКазы Р имеются последовательности, комплементарные инвариантным районам молекул тРНК. В процессе реакции несомненно происходят взаимные конформационные структурные перестройки M1 РНК и тРНК, которые значительно легче протекают в присутствии белка, однако детали этого процесса не выяснены. Мутантные тРНК с нарушенной вторичной или третичной структурой не подвергаются процессингу. Удастся осуществить реконструкцию активной РНКазы Р, если объединять белковый и РНК-компоненты, взятые из разных источников, например из клеток бактерий и человека. РНК-компоненты РНКазы Р разного происхождения не обладают гомологи-

ей по нуклеотидной последовательности, хотя отдельные короткие участки молекулы могут сохранять большое сходство. Следовательно, за узнавание РНК другой РНК могут отвечать как общая конформационная структура РНК, так и конкретные нуклеотидные последовательности. Вероятно, в процессе эволюции отбор был направлен на поддержание тех функциональных свойств субъединиц РНКазы Р, которые обеспечивают сохранность структур высшего порядка.

РНКазы Р представляет собой удивительную модель, пригодную для выяснения закономерностей белково-нуклеинового узнавания и позволяющую исследовать, каким образом белок, необходимый для работы фермента в клетке, может облегчать катализ, осуществляемый РНК.

РНК в составе РНКазы Р выступает как настоящий катализатор, свойства которого определяются нативной структурой РНК. Отмечают общность характеристик каталитических процессов, осуществляемых белками и РНК. В обоих случаях катализ основан на способности образовывать специфические комплексы с субстратом, характеризуется высокой субстратной специфичностью, а кинетика процесса описывается уравнением Михаэлиса — Ментен. Молекулы РНК, способные осуществлять биокатализ, называют *рибозимами*. Рибозимы отличаются от белков-катализаторов значительно меньшим разнообразием боковых радикалов. Вероятно, как это предвидел Крик, молекулы РНК на заре жизни представляли собой простейшие катализаторы (автокатализаторы), которые способствовали возникновению разнообразия нуклеотидных последовательностей. Скорее всего, именно они были способны к пребиотической эволюции, а затем сыграли большую роль в эволюции живых систем. Однако рибозимы — это не «молекулярные ископаемые» в ряду биологических катализаторов, они сохранились и активно функционируют в клетках про- и эукариот до наших дней.

Особенности процессинга тРНК у эукариот обусловлены тем, что ядерные предшественники многих тРНК содержат интрон. Сразу за антикодоном (от 5'-конца тРНК) следует нуклеотид, вслед за которым начинается последовательность одного интрона, обычно включающая 14—16 нуклеотидов (рис. 100, а). Созревание тРНК требует удаления интрона и сплайсинга. Сплайсинг предшественника тРНК в отличие от сплайсинга ядерных предшественников рРНК у инфузорий целиком зависит от присутствия белков-ферментов. Механизм реакций сплайсинга предшественником тРНК с участием ферментов, очищенных из дрожжей, изображен на рис. 100, б. Эндонуклеаза расщепляет РНК в двух сайтах сплайсинга, образуя на концах сшиваемых экзонов 2'—3'-циклофосфат и 5'-ОН-группу. Эндонуклеазная активность связана с ядерной мембраной. Это обстоятельство указывает на сопряженность процессинга с транспортом тРНК из ядра в цитоплазму. Следующие химические реакции (киназная реакция, в результате которой фосфорилируется 5'-

В следующем разделе будет рассмотрен механизм сплайсинга предшественников информационной РНК (мРНК), который осуществляется другими путями.

4. ПРОЦЕССИНГ РНК, СИНТЕЗИРУЕМЫЙ С ПОМОЩЬЮ РНК-ПОЛИМЕРАЗЫ II, У ЭУКАРИОТ

Модификация 5'-конца РНК и сплайсинг

РНК полимеразы II транскрибируют все гены эукариот, кодирующие белки, а также *малые ядерные РНК* и, по-видимому, РНК-компонент РНКазы Р. Рассмотрим лишь наиболее изученный механизм процессинга транскриптов, приводящий к образованию зрелых молекул мРНК, обычно содержащих 1500—2000 нуклеотидов.

Протяженность первичных ядерных транскриптов, образуемых РНК-полимеразой II, сильно варьирует, но может достигать десятков тысяч нуклеотидных пар, т. е. соответствует размерам ряда эукариотических генов (см. гл. IX). При исследовании клеточного ядра специальными методами электронной микроскопии удается обнаружить транскрипционные комплексы (рис. 101). Продвижение РНК-полимеразы по ДНК сопровождается образованием транскриптов, которые, взаимодействуя с белками, упаковываются в рибонуклеопротеидные комплексы.

Протяженные транскрипты эукариотических генов содержат последовательности интронов (см. гл. IX, раздел 2), которые при образовании мРНК вырезаются, тогда как нуклеотидные последовательности экзонов сшиваются, т. е. происходит процесс сплайсинга (рис. 102).

Разорванные (мозаичные) гены, составленные из экзонов и интронов, характерны для самых разных представителей эукариот — растений, дрожжей, различных беспозвоночных (черви, насекомые), птиц и млекопитающих, включая человека. Наличие и расположе-

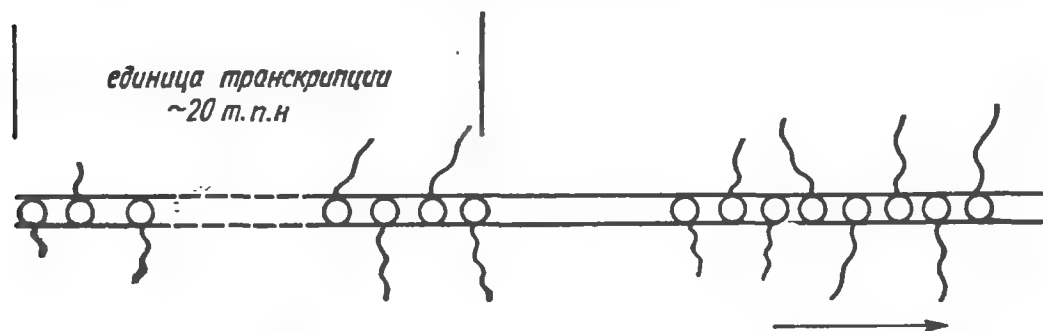


Рис. 101. Схема электронно-микроскопической картины транскрипции эукариотических генов

Стрелка указывает направление транскрипции; ветви «елочек» — новообразованные РНК в составе рибонуклеопротеидных комплексов

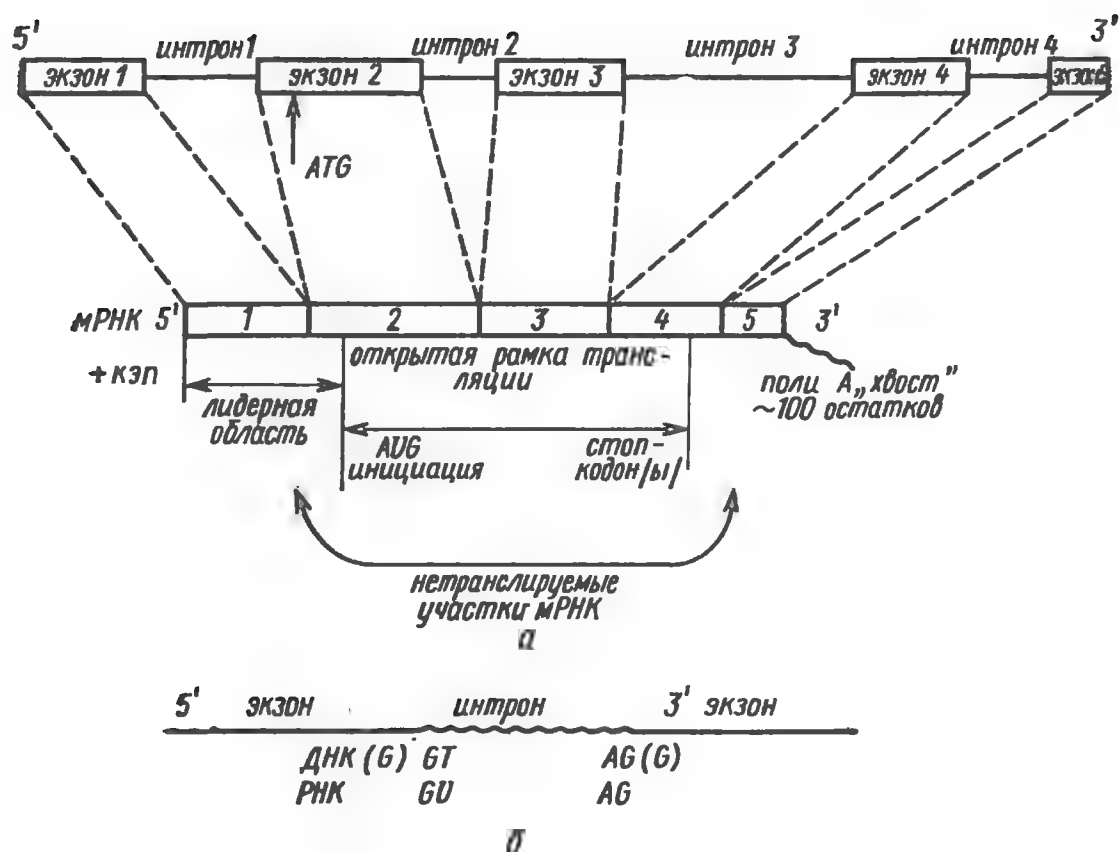


Рис. 102. Схема сплайсинга экзонов при образовании зрелой мРНК (а), границы экзонов и интронов (правило сплайсинга) (б)

ние интронов удастся наглядно показать при исследовании гибридов клонированных фрагментов геномной ДНК и зрелых мРНК, Участки геномной ДНК, включающие интроны, образуют петли, поскольку они не гибридизуются с мРНК. На рис. 103 демонстрируется структура таких гибридов мРНК и ДНК в случае гена α -коллагена цыпленка.

Число интронов может сильно варьировать: от одного в гене актина дрожжей до нескольких десятков (17 в гене δ -кристаллина птиц, 31 в гене вителлогенина птиц и 50 в гене проколлагена млекопитающих). Интроны могут составлять большую часть мозаичного гена. Суммарная нуклеотидная длина интронов может превышать во много раз длину экзонов. Например, в гене проколлагена птиц экзоны составляют лишь восьмую часть общей длины гена, а в гене тиреоглобулина млекопитающих — лишь двадцатую. Экзонами называют не только те районы гена, которые содержат последовательности, кодирующие участки полипептидной цепи, но и районы, транскрипты которых входят в состав зрелых мРНК, но не транслируются. Эти районы соответствуют «лидерной» области перед иницирующим кодоном (см. рис. 102) или в нетранслируемой части на 3'-конце мРНК вслед за терминирующим триплетом. Иногда, как это показано на рис. 102, а, лидерная зона соответствует отдель-

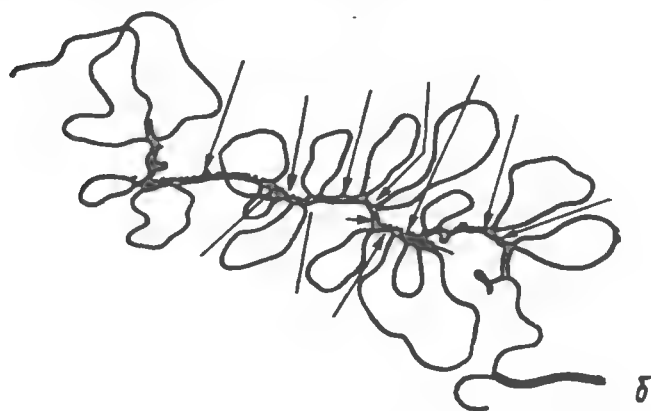
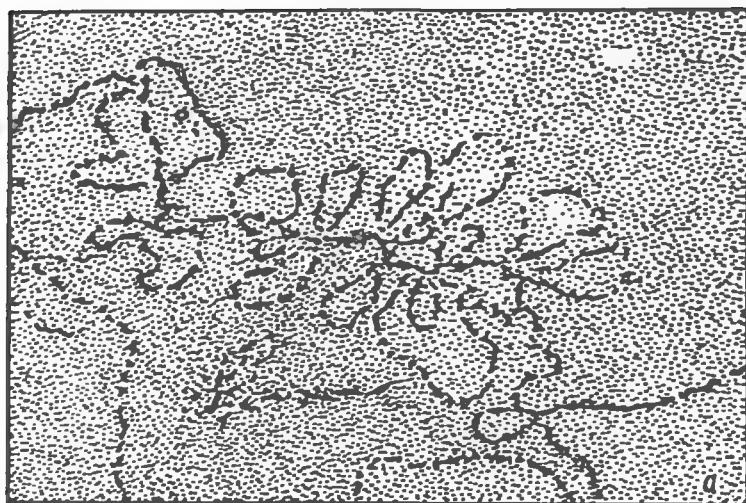


Рис. 103. Гибридные структуры мРНК и фрагмента гена $\alpha 2$ -коллагена цыпленка (электронная микрофотография)

Стрелки указывают районы РНК-ДНК гибрида (экзоны); петли содержат интроны

ным экзонам (ген овальбумина птиц или инсулина млекопитающих), 3'-нетранслируемый конец также может быть представлен отдельным экзоном (мРНК фибриногена крысы).

Инициация транскрипции сопровождается модификацией 5'-конца РНК с образованием специфической нуклеотидной структуры, в которой X7-метилированный остаток гуанозин-5'-трифосфата соединен 5'—5'-фосфодиэфирной связью с концевым нуклеотидом РНК. Это так называемый *кэп* (англ. cap), который присоединяется сразу после инициации транскрипции к адениловому или гуаниловому остаткам. Поэтому сайт инициации транскрипции часто называют *кэп-сайтом*.

Модифицированный 5'-конец обеспечивает эффективную трансляцию мРНК и удлиняет время ее жизни в клетке. Образование кэпа также способствует дальнейшему ходу процессинга. Узнавание кэпа, по-видимому, является важной стадией в процессе работы сложной ферментативной системы, осуществляющей сплайсинг.

Добавление структурных аналогов кэпа-динуклеотида — $m^7G(5)-p-p-p-5-N$ или m^7CTP — сильно подавляет сплайсинг.

Транскрипция экзонов и интронов сложного эукариотического гена осуществляется в порядке их расположения друг за другом (*коллинеарная транскрипция*), после чего вырезаются районы интронов, а экзоны сшиваются в результате сплайсинга. Порядок вырезания районов интронов необязательно соответствует последовательности их расположения в гене. Образуется мРНК, в которой вслед за лидерной зоной следует так называемая «открытая рамка трансляции» (англ. open reading frame) — последовательность триплетов, кодирующих полипептид, начинающаяся с иницилирующего AUG. Сплайсинг должен быть очень точным, в противном случае ошибка на один нуклеотид может вызвать нарушение «рамки» трансляции, что приводит, например, к образованию терминирующих триплетов и остановке синтеза полипептида. Последовательность ДНК в интроне начинается, как правило, с нуклеотидов GT с 5'-конца и заканчивается нуклеотидами AG (см. рис. 102, б). Соответственно в РНК последовательность интрона содержит по концам динуклеотиды GU и AG *. Однако границы экзонов и интронов часто не удается определить с точностью до нуклеотида, поскольку в районе сайтов сплайсинга нуклеотиды (например, G, см. рис. 102) могут быть повторены.

Нарушения сплайсинга могут быть губительны для организма. У человека известна аутосомная рецессивная мутация в гене фенилаланингидроксилазы. В гомозиготном состоянии мутация приводит к заболеванию — фенилкетонурии, проявляющейся в резкой задержке развития и смерти в раннем детском возрасте при отсутствии специальной диеты. Это наиболее частый случай врожденных ошибок аминокислотного метаболизма (1 : 8000 в США). Отсутствие фермента, катализирующего гидроксилирование фенилаланина с образованием тирозина, приводит к накоплению фенилаланина и нарушению образования многих метаболитов, образующихся из ароматических аминокислот. Протяженный ген фенилаланингидроксилазы человека, включающий около 90 т. п. н., содержит 12 интронов, составляющих основную часть гена. Единственная замена в 5'-районе 12-го интрона, превращающая канонический динуклеотид GT в AT, приводит к нарушению сплайсинга и процессинга. Образующаяся аберрантная РНК, не содержащая концевых экзонов, служит матрицей для трансляции преждевременно обрывающегося полипептида, который быстро распадается в клетке. Нарушения сплайсинга описаны также для генов человека, кодирующих глобины. Эти мутации у гомозигот приводят к болезням у человека, обусловленным недостаточностью того или иного типа глобина (гемоглобинопатии, или талассемии).

* Всегда приводится последовательность той или иной нити ДНК, которая идентична нуклеотидной последовательности транскрибируемой РНК, за исключением замены Т на U.

Ферментативная машина сплайсинга, локализованная в ядре, изучена недостаточно, однако ряд важных результатов получен при исследовании превращений предшественников индивидуальных мРНК (например, β -глобина) в ядерных экстрактах *in vitro*. Были выявлены следующие стадии сплайсинга (рис. 104). Прежде всего происходит разрыв нуклеотидной последовательности на границе экзона 1 и интрона с образованием 5'-конца, представленного фосфорилированным гуанозином — каноническим нуклеотидом левого (5'-конца) интрона. Следующая (или сопряженная с предыдущей) реакция заключается во взаимодействии 5'-конца удаляемой последовательности с 2'-гидроксилом аденозина интрона с образованием 2'—5'-фосфодиэфирной связи. Образуется разветвленная структура РНК, напоминающая лассо (кольцо с хвостом). Образование такой структуры можно также рассматривать как результат реакции трансэтерификации, в результате которой 2'-ОН-группа аденозина атакует фосфодиэфирную связь, замещая 3'-ОН-группу экзона 1. Теперь экзон 1 представлен отдельной молекулой РНК, удерживаемой вместе с другой молекулой, содержащей *лассо-интрон* и экзон 2, в едином комплексе с белками, участвующими в реакции. По-видимому, образование лассо является условием для завязывания ковалентной связи 3'-ОН-группы линейной молекулы экзона 1 с экзоном 2 в районе сайта сплайсинга; эта реакция также представляет собой трансэтерификацию, когда 3'-ОН-группа экзона 1 атакует фосфодиэфирную связь и замещает 3'-ОН-группу интрона; последовательность интрона выщепляется из РНК в форме лассо. В результате число фосфодиэфирных связей в составе субстрата и продукта не изменяется. Вероятно, процесс может идти без гидролиза высокоэнергетических кофакторов, как и аутосплайсинг предшественника рРНК.

По-видимому, в образовании лассо участвует последовательность сайта «ветвления» РНК, окружающая участвующий в реакции аденозин и способная образовывать комплементарную структуру с 5'-концом интрона (рис. 104). Внутренняя часть интрона, в ряде случаев достаточно протяженная, может быть безболезненно удалена без нарушения сплайсинга. Вопрос о том, какова судьба и возможная роль выщепляемых интронов, остается не ясным.

Необходимым компонентом системы сплайсинга гигантских ядерных предшественников мРНК являются так называемые малые ядерные РНК. Эти РНК обогащены уридином, поэтому они получили название U РНК: U1, U2, U3, U4 и т. д. Они легко разделяются с помощью электрофореза. Разные малые ядерные РНК отличаются числом нуклеотидов, входящих в их состав (от 90 до 400). Обнаружена исключительная консервативность нуклеотидных последовательностей малых ядерных РНК птиц, млекопитающих и дрозофилы.

Малые ядерные РНК имеют на 5'-конце специфический кэп, представленный триметилгуанозином: 5'-m³GpppAUACUUA... . Эта

структура в составе УРНК необходима для осуществления сплайсинга. Удаление кэпа и прилегающих к нему нуклеотидов из УРНК, равно как и добавление антител против триметилированного гуанозинового остатка, останавливает сплайсинг.

Вначале предполагали, что участие малых ядерных РНК в сплайсинге определяется тем, что их нуклеотидная последовательность комплементарна граничным экзон-интронным нуклеотидным последовательностям. Подобные комплементарные взаимодействия могли бы обеспечить «стягивание» экзонов и способствовать сплайсингу. Эта модель сплайсинга не получила экспериментального подтверждения, однако комплементарные взаимодействия отдельных УРНК с 5'-участками интрона (рис. 104, б), с 3'-сайтом сплайсинга или с районом ветвления РНК при образовании лассо были доказаны. Несомненно, они играют существенную роль при протекании реакций сплайсинга.

Структурная роль малых ядерных РНК в сплайсинге остается невыясненной, но ясно одно, что они являются важным и необходимым компонентом реакций сплайсинга. Можно предполагать, что отсутствие того или иного типа УРНК будет приводить к нарушению процессинга и, следовательно, сопровождаться прекращением экспрессии гена. УРНК кодируются десятками генов, обычно рассеянных по геному. Образование зрелых УРНК, синтезируемых с участием РНК-полимеразы II, также сопряжено с процессингом предшественника. Показано, что синтез отдельных УРНК приурочен к определенным стадиям развития амфибий (*Xenopus*). Дифференциальная экспрессия генов разных типов УРНК может определять регуляцию сплайсинга транскриптов генов в процессе развития организма.

Процессинг предшественников мРНК в ядре несомненно идет при участии белков. Гигантские ядерные транскрипты связываются с ядерными белками. Возникают структуры типа «бусин на нити», соединенных участками РНК (сравни с нуклеосомами, гл. XII). После обработки РНКазой эти структуры распадаются с образованием отдельных рибонуклеопротеидных частиц. Каждая частица включает около 500 нуклеотидов и по крайней мере 8—10 разных белков. Участие ряда белков в сплайсинге прямо доказано: моноклональные антитела, специфически реагирующие с отдельными белками, подавляют сплайсинг *in vitro*. Ядерные УРНК также связаны со специфическими белками, образуя малые ядерные рибонуклеотидные частицы. Способ функционирования этих частиц при сплайсинге не ясен, а белки, образующие эти частицы, охарактеризованы недостаточно.

Упаковка длинных ядерных транскриптов в составе рибонуклеопротеидных комплексов, по-видимому, необходима для осуществления правильного процессинга и узнавания сигналов сплайсинга. Элементом такой упаковки служит *сплайсосома* (англ. spliceosome), включающая УРНК, белки и субстраты сплайсинга: 5'-экзон, струк-

туру лассо-интрон и 3'-экзон. Молекулярная масса сплайсосомы около $3 \cdot 10^6$. Образование сложного комплекса сплайсосомы *in vitro* требует АТР. Негидролизуемые аналоги не заменяют АТР; следовательно, при образовании комплекса в отличие от химических реакций сплайсинга затрачивается энергия. Структура сплайсосомы и ее компоненты еще плохо изучены.

Удаление интронов, по-видимому, не идет строго и последовательно в направлении $5' \longrightarrow 3'$. Вероятно, в результате вырезания интрона из предшественника меняется конформация гигантской РНК, направляются и облегчаются отдельные этапы сплайсинга. Перспективным в познании закономерностей сплайсинга является создание искусственных генов с заданными последовательностями интронов и исследование *in vitro* сплайсинга сложных транскриптов таких генов.

Процессинг 3'-конца транскрипта

На 3'-конце большинства эукариотических мРНК (исключения составляют мРНК гистонов) находится последовательность из остатков адениловой кислоты, обычно включающая около 100 нуклеотидов. Считается, что присоединение адениловых остатков (полиаденилирование) приводит в первую очередь к увеличению времени жизни мРНК, ее стабилизации, а также обеспечивает транспорт мРНК в цитоплазму. Рассмотрим процессинг 3'-конца транскриптов, который сопряжен со сплайсингом или даже предшествует отдельным стадиям сплайсинга (рис. 105, а).

Транскрипция гена с помощью РНК-полимеразы II, как это было показано, например, для гена β -глобина млекопитающих и овальбумина птиц, может продолжаться еще на сотни нуклеотидов от сайта, соответствующего 3'-концу зрелой мРНК. Сигнальные нуклеотидные последовательности на 3'-конце гена и факторы, определяющие терминацию транскрипции с помощью РНК-полимеразы II, недостаточно исследованы. 3'-конец зрелой молекулы РНК определяется специфической эндонуклеазой, разрезающей длинный транскрипт. Эта реакция протекает быстро, отрезаемый эндонуклеазой дистальный конец мгновенно деградирует. Положение 3'-конца зрелой мРНК, к которому прикрепляются адениловые остатки, определяется последовательностью ААТААА на 3'-конце гена. Эта последовательность в составе зрелой мРНК (ААУААА) обнаруживается отступя 11—30 нуклеотидов (в среднем 15) от 3'-конца. Эндонуклеаза, узнавая последовательность ААУААА (сигнала полиаденилирования), расщепляет транскрипт, а образующийся 3'-конец подвергается полиаденилированию. Сигнал полиаденилирования необходим, но не достаточен для эндонуклеолитической реакции на 3'-конце транскрипта. В процессинге 3'-конца мРНК участвуют последова-

тельности, которые не входят в состав зрелой мРНК и локализованы за сайтом эндонуклеазной атаки, ближе к 3'-концу транскрипта.

В составе транскрипционной единицы в ряде случаев (например, ген овальбумина птиц) обнаруживается более одного сигнала полиаденилирования. Сигналы полиаденилирования могут находиться в разных участках гена. Выбор одного из возможных сигналов полиаденилирования, вслед за которым следует эндонуклеазное расщепление транскрипта, может также рассматриваться как способ регуляции экспрессии гена, обеспечивающий нужное направление 3'-процессинга с образованием мРНК данного типа. Выбор разных сайтов полиаденилирования, имеющий регуляторное значение, осуществляется, например, при транскрипции гена, кодирующего образование двух белков: кальцитонина — гормона С-клеток щитовидной железы, регулирующего концентрацию кальция, и кальцитонинподобного биологически активного пептида в нервной ткани (рис. 105, б). Информационная РНК, кодирующая кальцитонин, образуется путем сплайсинга первых четырех экзонов гена, причем последний экзон содержит сигнал полиаденилирования, узнаваемый в клетках щитовидной железы. В клетках мозга образуется другой тип мРНК, включающей последовательности не только первых трех экзонов, но и пятого, и шестого, причем последний несет второй сайт полиаденилирования, узнаваемый в клетках нервной ткани. Функции белка, кодируемого этим типом мРНК, не известны. Несомненно, выбор одного из возможных сайтов полиаденилирования играет роль в тканеспецифической экспрессии генов.

Значимость сайтов полиаденилирования при экспрессии генов выявляется при молекулярном анализе природы мутаций. Оказалось, что один из случаев нарушения синтеза β -глобина человека обусловлен мутацией, которая привела к образованию лишнего сайта полиаденилирования и нарушила созревание нормальной мРНК.

Модификация 3'-конца транскрипта также может происходить с участием малых ядерных РНК определенного типа. Такой способ процессинга 3'-конца осуществляется, например, при образовании мРНК гистоновых генов, лишенных полиадеинового хвоста на 3'-конце. В ходе процессинга предшественника мРНК гистона H3 морского ежа принимает участие U7РНК. Разные гистоновые мРНК морского ежа содержат консервативную последовательность из 12 нуклеотидов на 3'-конце, представляющую собой палиндром (изображена в составе зрелой мРНК, рис. 105, в). Можно представить, что U7РНК за счет комплементарных взаимодействий раскрывает палиндромную шпильку с образованием одонитевой петли из 15 нуклеотидов, облегчающей действие эндонуклеазы. Показано, что последовательность CAAGAAAGA, взаимодействующая с U7РНК, но не входящая в состав зрелой мРНК, необходима для процессинга. Процессинг 3'-конца, идущий с участием малых РНК, ингибируется антителами, специфически реагирующими с рибонуклеопротеидными частицами, в состав которых входят эти РНК.

5. РЕГУЛЯЦИЯ ЭКСПРЕССИИ ГЕНОВ ПУТЕМ АЛЬТЕРНАТИВНОГО СПЛАЙСИНГА

Принципиальной является возможность образования нескольких разных типов мРНК в результате изменения хода сплайсинга одного и того же первичного транскрипта. Для разных генов показаны так называемые альтернативные пути сплайсинга, основанные на использовании разных экзонов одного гена при образовании мРНК.

В результате альтернативного сплайсинга зрелые молекулы мРНК, образуемые при транскрипции одного гена, включающего несколько экзонов, будут различаться набором экзонов, кодирующих отдельные участки молекулы белка. Кроме того, последовательность экзона в ходе одного пути сплайсинга может служить интроном в ходе альтернативного пути сплайсинга. Таким образом, разные способы экспрессии одного гена могут приводить к образованию

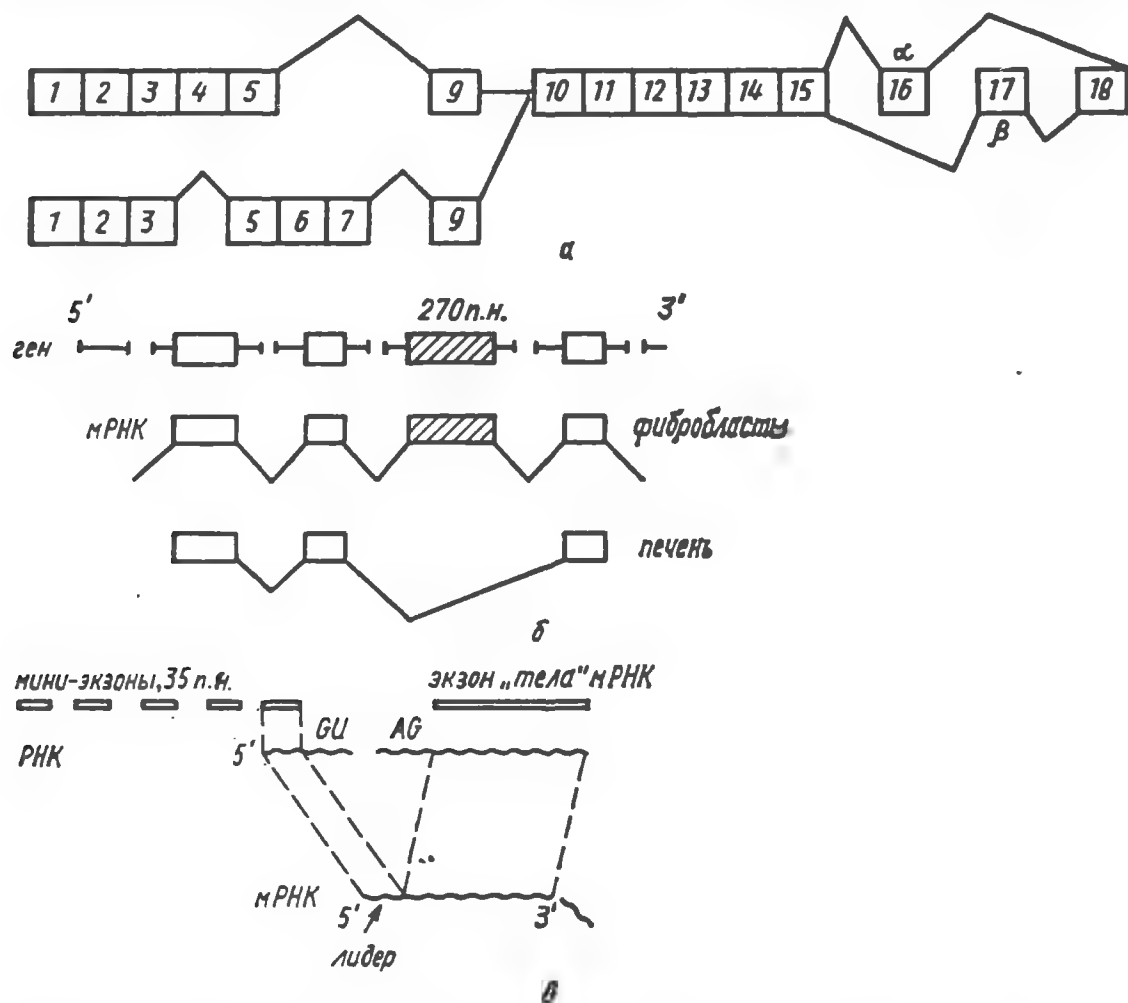


Рис. 106. Случаи альтернативного сплайсинга при созревании мРНК:

а — варианты сплайсинга экзонов гена тропонина, включая случай взаимноисключающей экспрессии экзонов 16 и 17 с образованием мРНК для α - и β -тропонина; прямоугольники — экзоны; ломаные линии — сплайсинг интронов; *б* — альтернативный сплайсинг при образовании фибронектина у человека (экзон, включающий 270 п. н., вероятно, кодирует участок полипептидной цепи фибронектина, отвечающий за связывание с поверхностью клетки); *в* — образование мРНК тубулинов трипанозом путем транс-сплайсинга (мини-экзоны и экзон «тела» мРНК расположены на разных хромосомах)

разных полипептидов (так называемых *изотипов* или *изоформ*). Например, это имеет место при экспрессии гена тропонина млекопитающих, включающего 18 небольших экзонов и кодирующего многочисленные изоформы этого мышечного белка. Разные типы тропонинов образуются в разных тканях на определенных стадиях развития.

Наличие путей альтернативного сплайсинга существенно увеличивает число разных мРНК, транскрибируемых с одного гена (рис. 106, а, б). При образовании мРНК тропонина с помощью механизма альтернативного сплайсинга используется также взаимоисключающая и взаимозаменяемая экспрессия экзонов 16 и 17, кодирующих определенный участок полипептидной цепи тропонина. На разных стадиях развития образуются α - и β -тропонины, различающиеся последовательностью из 14 аминокислот, начиная с 229-го и кончая 242-м аминокислотным остатком. Остальные участки полипептидной цепи этих изотипов тропонина идентичны. Остается не ясным, какие изменения функциональных свойств тропонина обусловлены экспрессией того или иного экзона в составе мРНК.

Функциональная роль отдельных экзонов при рассмотрении случаев альтернативного сплайсинга, возможно, прояснится на примере гена позвоночных, кодирующего полипептидные компоненты целой серии гликопротеидов — фибронектинов, секретируемых клеткой. Некоторые типы фибронектинов, являясь компонентами внеклеточного матрикса, связываются с клеткой и определяют свойства ее поверхности, другие находятся в плазме крови. Разные типы фибронектинов образуются путем альтернативного сплайсинга. Фибронектин плазмы, который не связан с клеточной поверхностью, синтезируется на мРНК, не содержащей одного из экзонов, возможно как раз того, который кодирует участок молекулы белка, отвечающий за связывание с клеткой.

Экспрессия самых разных генов может регулироваться путем выбора альтернативных путей сплайсинга. Например, явление альтернативного сплайсинга обнаружено при экспрессии гена, кодирующего основной белок миелиновых мембран, окружающих аксон и обеспечивающих эффективное проведение сигнала на большие расстояния. В результате сплайсинга синтезируются четыре формы основного белка миелина, специальные функции которых пока не исследованы. Альтернативный сплайсинг обеспечивает также разные пути экспрессии генов, кодирующих полипептидные гормоны, белки ионных каналов клетки, а также ядерные белки, участвующие в регуляции действия генов, определяющих ключевые стадии развития.

Альтернативный сплайсинг несет определенные выгоды, придавая своеобразную гибкость экспрессии эукариотических генов. Он создает разнообразие продуктов, кодируемых одним отрезком ДНК. По мере развития один и тот же отрезок ДНК используется для разных, хотя и сходных, целей. Ранее предполагали, что такой принцип экспрессии генов используется компактными геномами

вирусов, где необходима экономия информационного содержания генов. Теперь стало очевидным широкое распространение такого способа регуляции экспрессии генов. Неизвестно, каким образом регулируются пути альтернативного сплайсинга. Информация, определяющая потенциальные пути сплайсинга, вероятно, в определенной степени заложена в структуре предшественника мРНК, создающейся в результате комплементарных взаимодействий. Вероятными кандидатами на регуляторную роль являются малые ядерные РНК, меняющие характер этих комплементарных взаимодействий, а также белковые факторы, природа которых остается невыясненной. Те и другие могут рассматриваться как транс-действующие факторы, кодируемые генами, регулируемыми экспрессию других генов.

6. ТРАНС-СПЛАЙСИНГ

Сплайсинг рассматривался на примерах превращения одной молекулы РНК. Однако принципиальный интерес представляет обнаружение случаев *транс-сплайсинга*, когда происходит образование ковалентных связей между фрагментами РНК, синтезированными на разных генах. Эти гены могут быть локализованы даже в разных хромосомах. Вероятно, этот процесс облегчается в результате установления временных комплементарных взаимодействий между разными молекулами РНК. Транс-сплайсинг хорошо показан у паразитических одноклеточных эукариот — трипанозом. Ряд мРНК трипанозом, например для белков цитоскелета — тубулинов, образуется путем сплайсинга фрагментов РНК, транскрибированных с «мини-экзонов», включающих по 35 п. н., с экзоном «тела» мРНК (рис. 106, в). Мини-экзоны кодируют 5'-лидерную часть мРНК. Мини-экзоны и экзоны тела мРНК могут располагаться в разных хромосомах. Транскрипты мини-экзонов содержат два канонических нуклеотида GU, а транскрипт экзона тела — AG-пару. Следовательно, в основе транс-сплайсинга лежат те же молекулярные механизмы, которые функционируют в случае внутримолекулярного сплайсинга. Существование транс-сплайсинга и альтернативного сплайсинга сильно расширяет возможности образования регулируемой и достаточно разнообразной популяции клеточных РНК, кодируемых ограниченным числом генов.

ЛИТЕРАТУРА

- Кавсан В. М. Сплайсинг 2. Сплайсинг мРНК в ядре//Молек. биол. 1986. Т. 20. Вып. 6. С. 1451—1471.
- Чек Т. РНК как фермент//В мире науки. 1987. № 1. С. 26—36.
- Processing of RNA. CRC Press, ed. by D. Apirion, 1984.
- Padgett R. A., Grabowski P. J., M. M. Kanarska et al. Splicing of messenger RNA precursors//Ann. Rev. Biochem. 1986. 55. P. 1119—1150.
- Breitbart R. E., Andreadis A., Nadal-Ginord B. Alternative splicing: a ubiquitous mechanism for the generation of multiple protein isoforms from single genes//Ann. Rev. Biochem. 1987. 56. P. 467—495.

Глава IX

ГЕНОМ ЭУКАРИОТ

Геном эукариот обеспечивает сложнейшие программы развития и клеточной дифференцировки, которые осуществляются в результате последовательной активации и инактивации множества генов, взаимодействующих друг с другом. Эукариотическая клетка содержит во много раз больше генов, чем прокариотическая. Ниже приведено содержание ДНК в разных организмах (п. н. в расчете на гаплоидный геном):

<i>E. coli</i>	$5 \cdot 10^6$
Дрожжи <i>Saccharomyces cerevisiae</i>	$5 \cdot 10^7$
<i>Drosophila melanogaster</i>	10^8
Нематода <i>Caenorhabditis elegans</i>	10^8
Амфибии	$1-5 \cdot 10^9$
Высшие растения (сильно варьирует) <i>Lilium longiflorum</i>	$1,8 \cdot 10^{11}$
Млекопитающие	$3 \cdot 10^9$

В ядрах клеток дрожжей, насекомых, червей содержится в 5—10 раз, а у млекопитающих в несколько сотен раз больше ДНК, чем в клетке *E. coli*. Содержание ДНК в расчете на гаплоидный геном в целом увеличивается с возрастанием сложности организма. У амфибий и растений оно сильно варьирует от вида к виду и может значительно (в 10 раз и более) превышать количество ДНК в клетках млекопитающих. Однако было бы неверным считать, что прогрессивная эволюция, как правило, сопровождается увеличением содержания ДНК в расчете на гаплоидный геном. Известны также случаи, когда достаточно близкие виды содержат количество ДНК, различающееся в несколько раз. Это явление описано как парадокс содержания ДНК (англ. C value paradox), который до сих пор не получил достаточно определенного объяснения. Таким образом, размеры геномов не коррелируют с тем количеством ДНК, которое предназначено для выполнения функции кодирования белков.

Содержание ДНК в расчете на клетку обычно сохраняется постоянным в разных тканях одного организма. Отклонения от этого правила редкие. К ним относятся случаи образования в некоторых типах клеток политенных (многонитчатых) хромосом, образующихся в результате многократной редупликации ДНК без расхождения двуспиральных молекул, а также классические примеры утери ДНК («диминуция хроматина») в соматических клетках. Потери участков хромосом, иногда достаточно крупных, составляющих существенную часть материала хромосомы, как правило, касаются гетерохроматических районов. Функциональная значимость образования политенных хромосом и случаев диминуций не ясна. Эти факты лишь подчеркивают правило постоянства содержания ДНК на клетку, которое отражает принцип дифференцировки, основан-

ный на избирательной активации или инактивации генов, а не на их физическом удалении из хромосом специализированных клеток.

Подавляющее количество ДНК сосредоточено в ядре, обычно лишь небольшая часть ДНК находится в составе генома цитоплазматических органелл. Митохондрии грибов и млекопитающих содержат менее 1 % всей ДНК, а пластиды растений — 1—10 %. В клетках дрожжей *Saccharomyces cerevisiae* количество митохондриальной ДНК может достигать 20 % от всей клеточной.

Генетики сделали первые грубые оценки числа генов в эукариотической клетке путем сравнения частоты возникновения мутаций в данном гене и в целой хромосоме (или во всем геноме). Вероятность возникновения мутаций по данному гену по сравнению с вероятностью мутирования всей совокупности генов данной хромосомы тем меньше, чем больше разных генов в ней содержится. Так, в X-хромосоме *Drosophila melanogaster* мутация по данному гену возникала в 1000 раз реже, чем мутации во всей X-хромосоме. Следовательно, X-хромосома дрозофилы может содержать около 1000 генов. В хромосоме 3 дрозофилы, которая по массе вдвое больше X-хромосомы, мутационный тест выявляет около 2000 генов. Расчеты показывают, что весь геном дрозофилы содержит около 5000 генов. Дальнейшие исследования генетиков в целом подтвердили эти результаты, полученные еще в 30-х годах. Содержание ДНК в гаплоидном геноме *D. melanogaster* составляет около 10^5 п. н. Простой расчет показывает, что на один ген приходится $2 \cdot 10^4$ п. н., т. е. примерно в 20 раз больше ДНК, чем необходимо для кодирования полипептида средней молекулярной массы. Этот результат лег в основу представлений об «избыточном содержании» ДНК у эукариот, которые породили в 60—70-х годах целый ряд гипотез, отводивших избыточной ДНК роль в эволюции генома или регуляции активности генов.

В результате исследования клонированных генов эукариот удалось показать, что «избыточность» содержания ДНК, по крайней мере частично, объясняется наличием внутренних некодирующих районов гена (интронов), суммарная длина которых может значительно (в несколько раз) превышать длину частей гена, кодирующих полипептид. Районы, представляющие собой регуляторные участки гена, включают небольшую часть ДНК, обычно не превышающую нескольких сотен нуклеотидных пар.

1. КИНЕТИКА РЕАССОЦИАЦИИ ДЕНАТУРИРОВАННОЙ ДНК

Некоторые формальные характеристики эукариотического генома могут быть получены с помощью методов гибридизации нуклеиновых кислот, разработанных в 60-х годах. Эти методы прежде всего позволяют получить суммарную валовую характеристику степени разнообразия последовательностей ДНК, образующих геном. Мож-

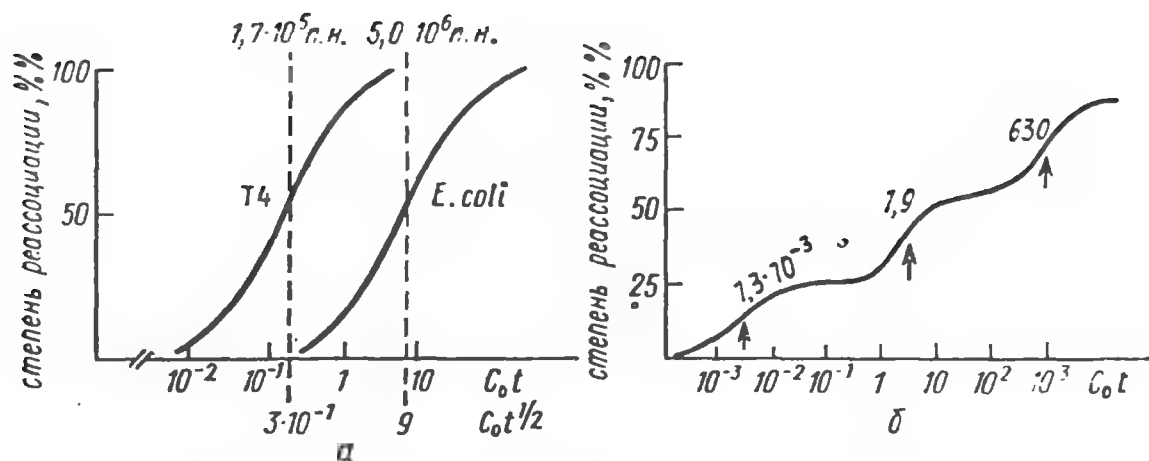


Рис. 107. Кинетика реассоциации денатурированных препаратов ДНК. (Значения C_0t (скорости реассоциации) отложены в логарифмическом масштабе.): а — значения $C_0t/2$ уменьшаются обратно пропорционально содержанию ДНК в геномах фага и бактерии; б — кинетика реассоциации ДНК тимуса теленка; стрелки указывают значения $C_0t/2$ для фракций генома.

но исследовать кинетику реассоциации одноплатных денатурированных фрагментов ДНК, определяя в каждый данный момент реакции количество ренатурировавшей двуспиральной ДНК. Кинетические параметры реакции количественно характеризуют степень разнообразия последовательностей ДНК, или сложность ДНК (рис. 107, а). Скорость реассоциации ДНК пропорциональна C_0t ($v = KC_0t$), где C_0 — молярная концентрация нуклеотидов ДНК; t — время, с. Значение $C_0t/2$, когда реассоциирует 50 % ДНК, характеризует степень сложности генома. Оказывается, что значение $C_0t/2$ для ДНК *E. coli* (~9) примерно в 30 раз больше, чем для ДНК фага Т4 ($3 \cdot 10^{-1}$). Действительно, содержание ДНК в клетках *E. coli* ($5 \cdot 10^6$ п. н.) также в 30 раз больше, чем в составе частицы фага ($1,7 \cdot 10^5$ п. н.). Сложность ДНК определяется общей длиной (в нуклеотидных парах) разных последовательностей ДНК, содержащихся в исследуемом образце ДНК. С увеличением степени полиплоидизации сложность геномной ДНК, как и кинетика реассоциации, меняться не будет.

При исследовании реассоциации ДНК эукариотической клетки выявляются «аномалии» кинетики реакции (рис. 107, б). Например, на кривой реассоциации ДНК из тимуса теленка можно различить ступеньки, которые отражают отдельные этапы реакции. Первая ступенька указывает на наличие быстро ренатурирующей фракции 1 (25 % ДНК). Следующая соответствует фракции 2, ренатурирующей со средней скоростью (30 % ДНК) и, наконец, остальная часть ДНК (45 %), представленная фракцией 3, ренатурирует медленно. Количество ДНК во фракции может быть определено исходя из значения $C_0t/2$. Так, например, значение $C_0t/2$ для фракции 2 (1,9) почти в 5 раз ниже ее значения для ДНК *E. coli*. В таком случае сложность фракции 2 составляет всего 10^6 п. н. В то же время коли-

чество ДНК во фракции 2, определенное химическими методами, составляет 10^8 п. н. Причина расхождения в 100 раз двух полученных величин состоит в том, что фракцию 2 в первом приближении можно рассматривать как состоящую из одинаковых повторяющихся последовательностей, представленных в геноме ста копиями. Еще более сильные расхождения между значениями количества ДНК, определяемого по $C_0t^{1/2}$ и химическими методами, получаются в отношении фракции 1. Другими словами, эта фракция состоит из нуклеотидных последовательностей, степень повторяемости которых в геноме особенно высока. Напротив, сложность фракции 3, выраженная числом нуклеотидных пар и соответствующая определенному значению $C_0t^{1/2}$, хорошо совпадает с результатом химического определения ДНК в этой фракции, составляющей около половины генома. Эта фракция состоит из последовательностей ДНК, представленных в геноме однократно (уникальные последовательности).

На рис. 107, б представлен лишь один из вариантов выявляемых кривых реассоциации ДНК эукариот, которые могут существенно различаться даже у близкородственных видов. Подобные результаты, полученные для геномов самых разных представителей эукариот, привели к заключению, что существенная часть генома (15—40 % или более, имеются значительные вариации для разных организмов) состоит из повторяющихся в разной степени последовательностей ДНК. Различают уникальные, умеренно повторяющиеся и высокоповторяющиеся последовательности ДНК эукариот, соответствующие трем выявляемым при ренатурации фракциям. Деление фракций на умеренно и высоко повторяющиеся достаточно условно. Гибридизационные методы дают достаточно грубую валовую характеристику генома в целом. Оценки числа копий генов, проведенные гибридизационными методами, не всегда давали правильные результаты вследствие ряда артефактов, сопровождающих эти эксперименты.

2. САТЕЛЛИТНЫЕ ДНК

Фракция ДНК, содержащая высокоповторяющиеся геномные последовательности, включает, по-видимому, функционально и в значительной степени структурно обособленную часть генома, представленную *сателлитными* ДНК. Сателлитные ДНК обладают характерным нуклеотидным составом и, следовательно, плотностью, отличающейся от валового нуклеотидного состава тотальной ДНК. Поэтому их удается в ряде случаев отделить от основной массы ДНК центрифугированием в градиенте плотности CsCl (рис. 108, а). Отдельные фракции сателлитных ДНК могут составлять до 10 % от общего содержания ДНК, как, например, один из сателлитов генома мыши. В составе одного генома можно обнаружить несколько разных сателлитных ДНК.

Определение нуклеотидных последовательностей сателлитных

ДНК показало наличие в их составе tandemных повторов из нескольких нуклеотидов. Главная повторяющаяся единица сателлитных ДНК (базовая последовательность), как, например, у *D. melanogaster*, может быть достаточно простой, состоящей из пяти или семи нуклеотидов: $(AATAT)_n$, $(AATAG)_n$, $(AATAC)_n$, $(AATAAAC)_n$ и $(AATAGAC)_n$. Встречаются и более сложные повторяющиеся единицы из 359 п. н. Базовая последовательность сателлитной ДНК повторяется многократно, на протяжении 10 тыс. п. н. и более. Таким образом, сателлитные ДНК образуют протяженные геномные блоки. Блоки, состоящие, например, из пяти и семи нуклеотидных tandemных повторов, могут в одной молекуле непосредственно прилегать друг к другу. Механизм возникновения таких структур не ясен. Нуклеотидные последовательности (234 п. н.) по длине сателлитной ДНК мыши могут варьировать за счет отдельных замен, вставок или делеций.

Меченые сателлитные ДНК можно гибридизировать с хромосомными препаратами *in situ*, а после гибридизации выявлять их локализацию. В хромосомах самых разных эукариот (дрозофила, млекопитающие) сателлитные ДНК преимущественно обнаруживаются в центромерных и теломерных районах хромосом (рис. 108, б). Это так называемые *гетерохроматические районы*, в которых материал хромосом сохраняется в процессе клеточного цикла в значительно более компактном состоянии в отличие от основной (эухроматической) части хромосом.

Наборы сателлитных ДНК могут сильно различаться у близких видов. Такая видовая специфичность рассматривалась как результат эволюционной нестабильности этого класса ДНК. Действительно, в процессе эволюции происходит амплификация одних видов сателлитных ДНК и диминуция других. В то же время отдельные сателлитные ДНК сохраняются идентичными, консервируются у видов, дивергировавших в эволюции более 50 млн. лет назад.

Ни одна из гипотез, предполагающих какую-либо функциональную роль сателлитных ДНК, пока не получила экспериментального подтверждения. Полагали, что сателлитная ДНК, локализуясь в определенных районах хромосом, может играть роль в узнавании

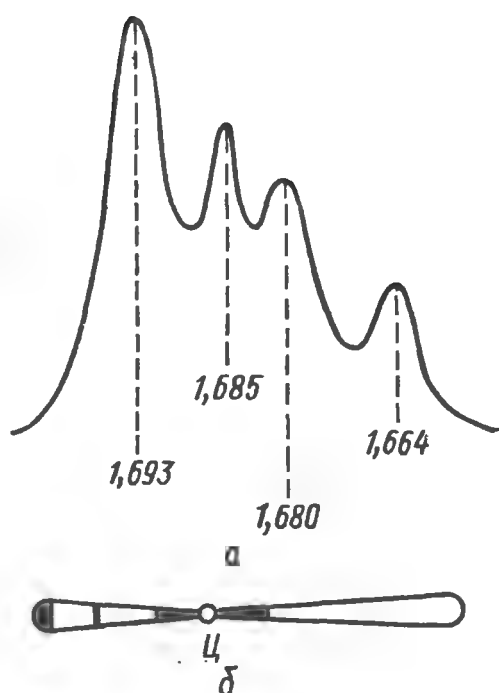


Рис. 108. Сателлитные ДНК: а — центрифугирование препарата ДНК *Drosophila virilis* в градиенте $CsCl$, основная масса ДНК имеет плавучую плотность 1,693; б — локализация сателлитных ДНК, выявляемая методом гибридизации *in situ* в хромосомах млекопитающих (схема); Ц — центромера

хромосом друг другом или их расхождении в процессе клеточных делений.

До сих пор, за некоторыми исключениями, не удавалось обнаружить транскрипции сателлитных ДНК.

3. УМЕРЕННО ПОВТОРЯЮЩИЕСЯ ПОСЛЕДОВАТЕЛЬНОСТИ

Обычно к этому классу относят последовательности, представленные в геноме десятками или сотнями копий. Среди умеренно повторяющихся последовательностей можно выделить две структурно и функционально различающиеся геномные фракции. Первая включает семейства генов, построенные из сгруппированных tandemно повторяющихся копий. Они участвуют в осуществлении жизненно важных общеклеточных функций. Это гены гистонов, транспортных РНК и рибосомных РНК. В целом они составляют несколько процентов от всей геномной ДНК. Число копий в составе одного семейства (например, генов гистонов) может варьировать от десяти до нескольких сотен у разных представителей эукариот.

Вторую большую разнородную группу составляют подвижные (мобильные) генетические элементы разной природы. На их долю приходится значительная часть генома — 10—20 %. В отличие от генов первой группы, осуществляющих общеклеточные функции, отдельные семейства которых сгруппированы в одном или нескольких районах эукариотического генома, подвижные элементы рассеяны по геному, перемежаясь с уникальными последовательностями ДНК (см. гл. XI).

В геноме обнаруживаются также рассеянные или находящиеся в кластерах гены, кодирующие гомологичные белки со сходными функциями. Например, это гены для разных типов актина, тубулина, белков оболочки яйца насекомых, гонадотропинов позвоночных. Перечисление этих генов показывает, что семейства таких генов могут выполнять как общеклеточные, так и специализированные функции. Такие *мультигенные семейства*, включающие по 5—20 копий и кодирующие белки со сходными функциями, также можно отнести к фракции умеренно повторяющихся последовательностей ДНК.

4. УНИКАЛЬНЫЕ ПОСЛЕДОВАТЕЛЬНОСТИ ГЕНОМА

Уникальные последовательности генома содержат не только гены, кодирующие белки, но и последовательности ДНК, расположенные между генами, а также в составе интронов, разделяющих участки ДНК, кодирующие полипептиды. Роль некодирующих уникальных последовательностей, составляющих основную часть эукариотического генома, остается до сих пор не выясненной.

Такова формальная валовая характеристика ДНК эукариот. Для понимания принципов структурно-функциональной организации генов необходимо было исследовать индивидуальные гены и разработать методы изучения их функционирования. Развитие работ в этом направлении связано с методологией, основанной на технике молекулярного клонирования. Удалось не только получить отдельные гены в количестве, достаточном для структурного анализа, но и найти эффективные способы оценки экспрессии изолированных генов. Особый интерес представляют возможности модификации клонированных генов с последующим определением их способности к функционированию в самых разных системах *in vivo* (в тканях трансгенных организмов) и *in vitro* (в культурах клеток, получивших ген, или в бесклеточных системах транскрипции). Такой подход используется так называемой суррогатной («обратной») генетикой; он отличается от методов классической генетики, в которых для исследования структуры гена и его регуляторных элементов отправной точкой служили прежде всего обнаруженные экспериментатором фенотипические изменения, обусловленные мутациями. Новая методология следует другой стратегии: разные участки клонированного гена с известной молекулярной структурой подвергают определенным модификациям (получают делеции, инверсии, нуклеотидные замены), после чего следят за изменившимся характером экспрессии гена в выбранной экспериментатором тест-системе. Таким образом можно выявить участки, отвечающие за регуляцию экспрессии гена.

Методы молекулярного клонирования открыли новую эру в исследовании эукариотических генов и выявили особенности их строения, которые ранее никто не мог предсказать. Общая схема строения эукариотического гена, содержащего экзоны и интроны, была представлена на рис. 102.

5. ПРЕРВАННЫЕ ГЕНЫ ЭУКАРИОТ (ЭКЗОНЫ И ИНТРОНЫ)

Существование интронов в эукариотических генах обеспечивает регуляцию экспрессии генов в развитии благодаря альтернативным путям сплайсинга, в основе которых лежит возможность использовать разные экзоны одного гена для образования разных мРНК. Кроме того, в интронах (т. е. внутри гена) могут находиться важные элементы регуляции транскрипции — усилители, или энхансеры (англ. enhancers) (см. гл. X).

Обнаружение интронов в большинстве эукариотических генов заставило задуматься о том, как и когда возникли интроны и какова их роль в эволюции генома.

Гены, разорванные интронами, возникли достаточно давно. Так, например, общий план строения генов β -глобина из двух интронов и трех экзонов очень сходен у всех животных, дивергиро-

вавших десятки миллионов лет назад. Более того, структура гена леггемоглобина растений, принимающего участие в регуляции содержания кислорода в клубеньковых бактериях, также укладывается в ту же общую схему строения глобинового гена. Положения отдельных интронов в генах триозофосфатизомеразы сходны у растений и позвоночных. Это позволяет предполагать, что интроны присутствовали в гене триозофосфатизомеразы уже 10^9 лет назад, а не образовались в результате более поздних независимых актов внедрений в одни и те же участки генов в процессе эволюции. На древнее происхождение интронов также указывает сходство механизма сплайсинга, основанного на узнавании сходных последовательностей на границах интронов у растений грибов, насекомых и позвоночных.

Было высказано предположение, что экзоны кодируют определенные автономные элементы укладки полипептидной цепи, представляющие собой функциональные сегменты белковой молекулы, которые сортируются в процессе эволюции. Если процессы такой перетасовки генетического материала, механизмы которых не рассматриваются, идут по районам интронов, то структура экзонов не изменяется и, следовательно, не нарушаются функциональные свойства отдельных белковых доменов. Экзоны могут соответствовать участкам доменов или отдельным белковым доменам, т. е. тем участкам белковой молекулы, которые можно выделить как пространственно делимые структуры, обладающие определенной биологической функцией. Установление размеров экзонов во многих генах показало, что главный класс экзонов имеет размеры около 140 п. н., что соответствует 40—50 а. о. в молекуле белка. Большая часть белковых доменов, содержащих в среднем 100—130 а. о., складывается из нескольких элементов вторичной структуры («супервторичных» структурных единиц), кодируемых отдельными экзонами. N-терминальный участок из нескольких гидрофобных аминокислот (сигнальный пептид) секреторных белков, как правило, также кодируется отдельным экзоном.

Если белок содержит ряд структурно сходных повторяющихся доменов, то наблюдается строгое соответствие отдельных экзонов доменам или субдоменам белковой молекулы. Гены, относящиеся к так называемому «сверхсемейству генов иммуноглобулинов», содержат разное число экзонов, кодирующих домены полипептидной цепи, каждый из которых включает около 110 а. о. Гомология между отдельными доменами этих белков, выполняющих разные функции в организме, наблюдается на уровне первичной, вторичной и третичной структуры. Гены этого семейства могут содержать один экзон (ген β_2 -микроглобулина), два или четыре (гены секретируемых антител В-клеток) и, наконец, пять экзонов (ген гликопротеина плазмы человека). β -Кристаллины мышцы содержат четыре белковых домена, каждый из которых включает определенный «структурный мотив» полипептидной цепи (рис. 109). Каждый домен

кодируется отдельным экзоном. В случае γ -кристаллинов кодируемая информация для двух «структурных мотивов» обнаруживается в составе одного экзона.

Примеры использования одного и того же экзона в разных генах подтверждают гипотезу о перетасовке экзонов в процессе эволюции. Ока-

зывается, что один и тот же структурный элемент может использоваться при образовании разных белков. Процесс перетасовки экзонов способен облегчить возникновение больших и сложных белков из небольших функциональных единиц — элементов супервторичной структуры (субдоменов или доменов). Ген одной из сериновых протеиназ (активатора плазминогена), участвующей в процессе свертывания крови, содержит 13 экзонов (рис. 110, а). Один из них кодирует структуру характерного белкового домена — «пальца». Этот экзон обнаруживается также в гене фибронектина, где он представлен несколькими копиями, разделенными интронами. Другой экзон гена плазминогена выявлен в гене эпидермального фактора роста, пять остальных экзонов формируют протеазный домен.

Ген рецептора липопроотеида низкой плотности, обеспечивающего транспорт холестерина, имеет размеры более 45 т. п. н. и содержит 18 экзонов, из которых часть также обнаружена в генах, кодирующих совсем другие функции (рис. 110, б). Рецептор является

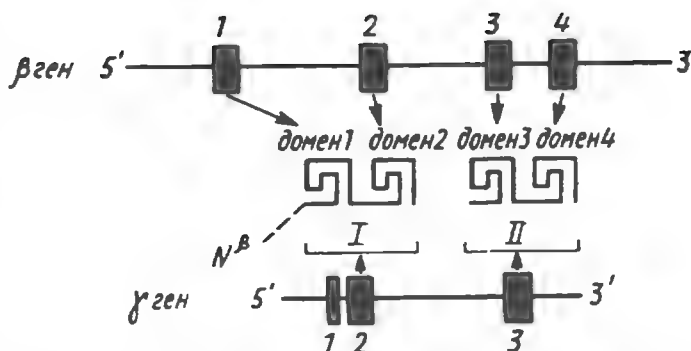


Рис. 109. Экзоны (черные прямоугольники) генов β - и γ -кристаллинов мыши и домены («греческий ключ») в составе полипептидной цепи

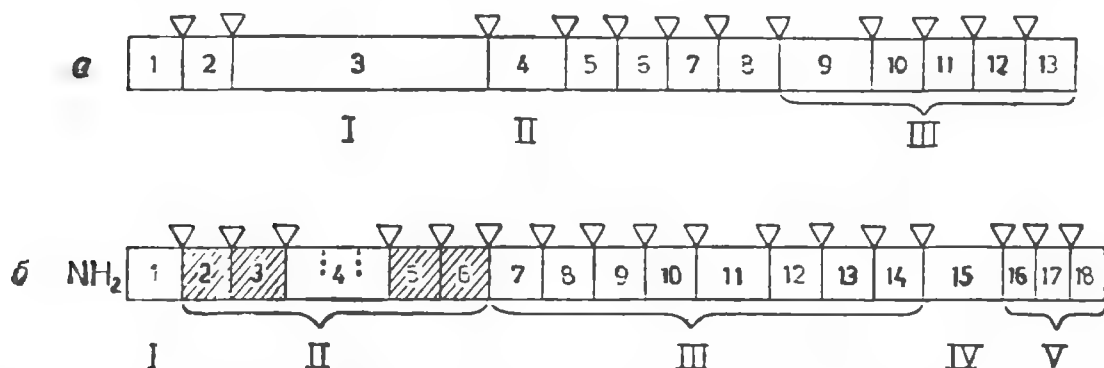


Рис. 110. Сложные структуры генов, вероятно, образованные в результате сортировки общих «предковых» экзонов

Прямоугольники — экзоны; треугольники — положение интронов; а — ген активатора плазминогена (сериновой протеиназы): I — «палец» фибронектина; II — гомология с экзонем эпидермального фактора роста; III — экзоны протеазного домена; б — ген рецептора липопроотеида низкой плотности: I — сигнальная последовательность; II — домен, связывающий липопротеид (гомология с экзонем в гене фактора комплемента C9 указана штрихами); III — гомология с экзонем гена эпидермального фактора роста; IV — домен, подвергающийся O-гликозилированию; V — кодируют трансмембранную и цитоплазматическую часть рецептора

мультифункциональным мембранным белком со связывающими липопротеид сайтами, обращенными наружу. Снаружи находятся также участки молекул, обеспечивающие эндоцитоз (поглощение клеткой) комплекса рецептора и липопротеида. Сигнальная пептидная последовательность белка, характерная для секреторируемых и мембранных белков, включена в состав первого экзона. Белковый домен, связывающий липопротеид, содержит семь повторяющихся элементов, кодируемых пятью экзонами.

По-видимому, два интрона утеряны. Полипептидная последовательность, кодируемая вторым, третьим, пятым и шестым экзонами, содержится также в составе фактора комплемента C9, где она также кодируется отдельными экзонами. Далее расположен район из восьми экзонов, он гомологичен району гена, кодирующему предшественник эпидермального фактора роста. Экзоны 7, 8 и 14 представляют собой повторы, кодирующие по 40 аминокислот и содержащиеся в генах, контролирующих процесс свертывания крови. Затем расположен экзон, кодирующий домен, обогащенный серином и треонином, который является мишенью О-гликозилирования рецептора. В итоге структура гена рецептора липопротеида низкой плотности в целом наглядно демонстрирует возможность перетасовки экзонов и соответствующих автономных функциональных структур сложной белковой молекулы.

Предполагается, что мозаичная экзон-интронная структура генов, свойственная эукариотам, вероятно, была более древней, чем безынтронная, прокариотическая. В таком случае традиционные филогенетические представления, согласно которым прокариот помещают в основание эволюционного древа, а эукариот — на вершины, должны быть пересмотрены. Геном прокариот, как правило, не содержащий генов с интронами, рассматривается как компактный (рационализированный), образовавшийся в результате потери интронов, например, в результате отбора на скорость репликации. Напротив, предполагается, что мозаичная структура генов определяет эволюционные возможности генома, тогда как прокариоты, утратившие интроны, представляют собой эволюционный тупик. Заметим, однако, что интроны, удаляемые в результате сплайсинга, изредка обнаруживаются при экспрессии генов в клетках бактерий, например в гене тимидилатсинтетазы фага T4.

Определение нуклеотидных последовательностей в составе гомологичных генов (например, генов глобинов), кодирующих полипептиды со сходным строением и функцией у одного или разных организмов, показало, что наибольшим изменениям в эволюции подвергались интроны, а не экзоны. В интронах обнаружены вставки, делеции и другие перестройки, в то время как последовательности экзонов оказываются значительно более консервативными. Изменения в нуклеотидных последовательностях экзонов часто обусловлены лишь отдельными нуклеотидными заменами. Эти наблюдения также можно истолковать в пользу представлений о том, что

перекомбинация экзонов осуществлялась по районам интронов, что сопровождалось изменением структуры последних.

Исследования ряда эволюционно связанных генов, содержащих гомологичные последовательности, позволили сделать заключение, что в эволюции происходили не только утери интронов, но и их приобретения. Механизм этого процесса не ясен. Возможно, вставки интронов происходили на уровне РНК. Если процесс вырезания интрона с помощью реакций трансэтерификации термодинамически обратим (см. гл. 8), то возможно и внедрение линейной молекулы в РНК с помощью реакции, обратной сплайсингу. На образовавшейся РНК как на матрице в результате обратной транскрипции может синтезироваться ДНК-копия, которая затем интегрируется в геном (см. гл. XI).

ЛИТЕРАТУРА

Б. Льюин. Гены.— М.: Мир, 1987.—544 с.

Эволюция генома. Сб./Под ред. Г. Довера и Р. Флавелла.— М.: Мир, 1986.—368 с.

Т. Г. Беридзе. Сателлитные ДНК.— М.: Наука, 1982.—120 с.

Gilbert W. Genes in pieces // *Nature*. V. 271. 501—504. 1978.

Gilbert W. Gene-in-pieces revisited // *Science*. 1985. V. 228. 823—824.

Blake C. F. Exons and the evolution of proteins // *Intern. Rev. of Cytol.* 1985. V. 93. P. 149—185.

Gilbert W. Marchiorne M., McKnight G. On the antiquity of introns // *Cell*. 1986. V. 46. 151—154.

Глава X

РЕГУЛЯТОРНЫЕ ЭЛЕМЕНТЫ ГЕНОВ И РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ У ЭУКАРИОТ

Инициация и регуляция транскрипции ДНК у эукариот с участием РНК-полимеразы в большей степени, чем у прокариот, зависит от множества других белков — факторов транскрипции, взаимодействующих с дискретными участками ДНК, образующих сложный эукариотический промотор. В районе промотора, прилегающего к сайту инициации транскрипции (*кэп*-сайту), обнаружены участки с характерными нуклеотидными последовательностями (мотивами), которые оказывают цис-действие на экспрессию близлежащего гена. Эти элементы могут взаимодействовать с РНК-полимеразой и другими белками-факторами транскрипции. Разные ядерные белковые факторы транскрипции, представляющие собой регуляторные белки, способны связываться с теми или иными нуклеотидными последовательностями ДНК, оказывая тем самым влияние на экспрессию разных генов. Такие белки, способные к диффузии

в ядре, являются транс-действующими факторами, определяющими эффективность экспрессии генов, расположенных в самых разных участках генома.

1. РЕГУЛЯТОРНЫЕ ЭЛЕМЕНТЫ ГЕНОВ, ТРАНСКРИБИРУЕМЫХ РНК-ПОЛИМЕРАЗОЙ II

Промоторные элементы генов одноклеточных эукариот — дрожжей — содержат сайты инициации (И), нуклеотидную последовательность ТАТА (обычно ТАТААА), а также другие элементы — активирующие последовательности (АП, UAS, англ. upstream activating sequences), находящиеся перед сайтом инициации транскрипции (рис. 111, а). Кроме того, промотор может содержать элементы оператора О, участвующего в репрессии транскрипции. Расстояние между ТАТА-элементом и сайтом инициации может варьировать от 40 до 120 п. н., и в отличие, например, от промоторов позвоночных в промоторах дрожжей правильная точная инициация транскрипции сохраняется при изменении расстояния между сайтом инициации и ТАТА-элементом. Инициаторный элемент представляет собой особый участок, включающий нуклеотидную последовательность TCG^A. Участок узнается РНК-полимеразой или связанными с ней белковыми факторами.

Элементы АП являются сайтами связывания белков, регулирующих транскрипцию. Ряд их свойств напоминает свойства усилителей транскрипции (энхансеров) в генах позвоночных, действующих вне зависимости от ориентации и на значительных расстояниях (около 1000 п. н.) от сайта инициации (см. раздел 2 этой главы).

Благодаря использованию большого набора мутаций по промоторам и генам активирующих белков дрожжей удалось выяснить некоторые особенности взаимодействия белков-активаторов с АП, а также характерные свойства этих белков. Белок GAL4 активирует гены, необходимые для катаболизма галактозы. GAL4 связывается с АП, представленной повторяющимися элементами по 17 п. н. Степень активирующего действия пропорциональна числу этих элементов в промоторе. Функция связывания ДНК и активации транскрипции принадлежит разным участкам белка GAL4, который содержит 881 аминокислоту. 73 остатка с N-конца молекулы белка достаточны для обеспечения специфического связывания с ДНК. Этот участок связывает ионы цинка и содержит характерную структуру — «цинковые пальцы», обнаруженные в целом ряде белков, активирующих транскрипцию (см. раздел 4 этой главы). Два других дискретных участка белка, включающих аминокислоты 149—196 и 768—881, достаточны для обеспечения активации транскрипции. Эти участки содержат кислые аминокислотные остатки. По-видимому, в разных активаторных белках эти районы обладают

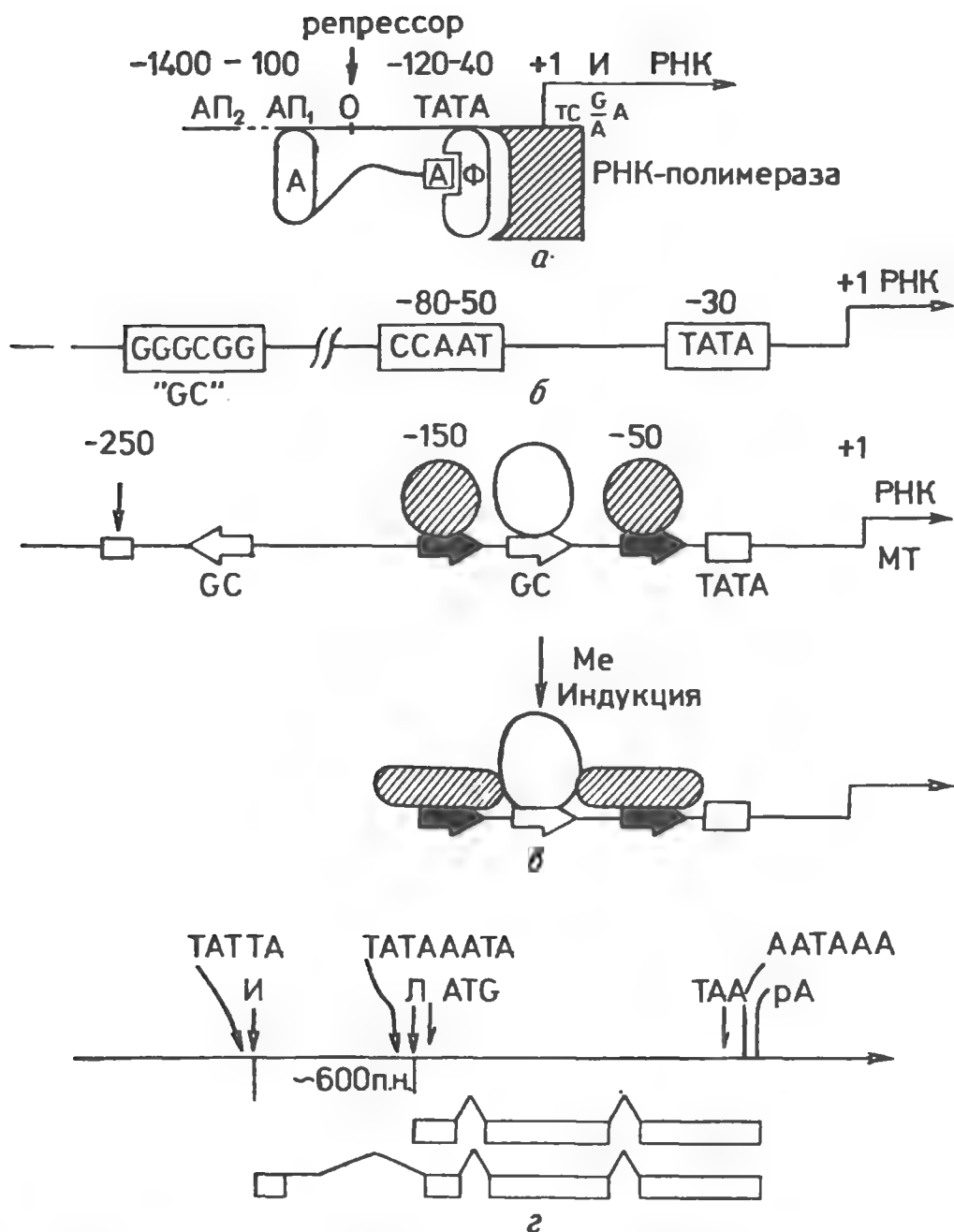


Рис. 111. Районы промоторов генов эукариот, транскрибируемых РНК-полимеразой II:

a — общая схема строения промотора генов дрожжей; АП — активирующие последовательности; А — участки молекулы белка-активатора; Ф — фактор транскрипции, связывающийся с ТАТА-элементом; И — сайт инициации; *б* — схема строения промотора гена у позвоночных; *в* — промотор гена металлотионеина (МТ) человека: черные стрелки — положение контролирующих элементов (Me), ответственных за регуляцию экспрессии генов тяжелыми металлами; кружки и эллипсы — белковые факторы и их предположительные взаимодействия при индукции транскрипции металлами; *г* — промоторы гена алкогольдегидрогеназы гена *D. melanogaster*; И, Л — сайты инициации транскрипции у имаго и личинки, им предшествуют «ТАТА-мотивы»; АТГ и ТАА-сайты инициации и терминации трансляции; прямоугольники — экзоны, ломаные линии — сплайсинг экзонов; рА — сайт полиаденилирования

сходными структурными особенностями, но не содержат гомологичных последовательностей. Предполагается, что эти белки стимулируют транскрипцию благодаря взаимодействию с другими фактора-

ми транскрипции, например с ТАТА-связывающимся белком, который в свою очередь взаимодействует с РНК-полимеразой (рис. 111, а). Белки — активаторы транскрипции, обладают, по крайней мере, двумя дискретными доменами, один из которых предназначен для узнавания ДНК, а другой для осуществления белок-белковых взаимодействий, играющих очень большую роль в образовании активного транскрипционного комплекса.

У высших многоклеточных эукариот (насекомых, позвоночных) в пределах 100—200 п. н. перед стартом транскрипции (рис. 111, б) была выявлена более сложная мозаика промоторных элементов, представленных короткими нуклеотидными последовательностями («мотивами»). На расстоянии 27—30 п. н. от *кэп*-сайта расположен ТАТА-мотив, усредненный вариант которого (так называемый *consensus*) можно представить как $\text{ТАТА}_T^A \text{А}_T^A$. Положение ТАТА-эле-

мента строго определяет сайт инициации транскрипции, т. е. 5'-конец транскрипта. При повреждении или удалении ТАТА элемента образуется набор молекул РНК с разными 5'-концами. Отдельные нуклеотидные замены в ТАТА-элементе могут приводить к резкому снижению эффективности транскрипции. Промоторная зона некоторых генов (например, генов яичного лизоцима птиц или гидроксиметилглутарилКоА-редуктазы — ключевого фермента биосинтеза холестерина у млекопитающих) не содержит ТАТА-элемента, и транскрипция начинается с нескольких разных сайтов. Образующиеся РНК различаются по 5'-концам в районе нетранслируемой лидерной (см. рис. 102) последовательности. Возможно, различные лидерные зоны определяют характер регуляции экспрессии генов на уровне трансляции. То же самое наблюдается и для генов дрожжей, лишенных ТАТА-элемента в районе промотора.

В пределах 100—200 п. н. перед стартом транскрипции многих генов находятся по крайней мере еще два коротких нуклеотидных «мотива», усредненные варианты которых можно представить как GGGCGG («GC-мотив») и ССААТ. Предполагается, что «GC-мотив», который может встречаться несколько раз по длине промоторной зоны, включающей 200—300 п. н., характерен для промоторных районов генов, работающих конститутивно и обеспечивающих обще-клеточные функции. Действительно, такие повторы обнаруживаются перед генами, кодирующими белки, необходимые для жизнедеятельности самых разных клеток организма. Перед геном дигидрофолат-редуктазы мыши вкраплены четыре таких мотива, перед генами глюкозо-6-фосфатдегидрогеназы и фосфоглицераткиназы человека соответственно девять и два. Нельзя исключить, что «GC-мотивы» принимают непосредственное участие и в индукции генной активности, поскольку они обнаружены и в промоторах ряда индуцируемых генов.

Последовательность ССААТ (СААТ) встречается в промоторной зоне разных тканеспецифических генов: в положении — 80—50 раз-

ных генов глобинов, в гене тиреоглобулина млекопитающих, в гене α -актина скелетных мышц цыпленка и других генах. Роль «мотива ССААТ» может быть достаточно важной в регуляции активности глобиновых генов, активируемых или репрессируемых на определенных стадиях развития. Репрессия синтеза γ -глобина плода у взрослого организма снимается в случае мутационного замещения одного нуклеотида в последовательности ССААТ. Мутация приводит к так называемой наследственной персистенции (сохранению синтеза γ -глобина плода у взрослых). В промоторах глобиновых генов отсутствует «GC-мотив».

Молекулярные механизмы, с помощью которых описанные элементы промотора регулируют транскрипцию, еще не выяснены, но несомненно, что активность промоторных элементов обусловлена связыванием с определенными белковыми факторами, обеспечивающими точную и эффективную транскрипцию генов РНК-полимеразой II. Выделены разные белки, взаимодействующие с разными участками промотора, содержащими ТАТА, ССААТ или «GC-мотив». По-видимому, существует несколько белков, способных связываться с «мотивом ССААТ», среди них — гетеродимер, состоящий из разных субъединиц. Белок, узнающий «GC-мотив», связывается с участком ДНК, включающим 18—20 п. н., в центре которого находится GC-элемент. Эффективность промотора, по крайней мере частично, определяется эффективностью отдельного элемента («мотива») в составе промотора, числом этих элементов и их взаимным расположением. Эти элементы, вероятно, функционируют в зависимости от ближайшего нуклеотидного окружения. Замены близлежащих нуклеотидов могут сильно сказываться на эффективности действия элемента. Так, например, замены выделенных жирным шрифтом нуклеотидов в окружении «GC-мотива» (GGGGCGGGGC) могут снижать активность промотора, тогда как замена первого G на T вполне допустима. Если область промотора содержит как GC, так и СААТ-элементы, то разные белковые факторы транскрипции, взаимодействующие с ними, могут согласованно активировать транскрипцию.

Гены, кодирующие адаптивные белки, образование которых резко усиливается под влиянием разных факторов среды (повышение температуры, отравление металлами), содержат в составе промоторов дополнительные характерные короткие нуклеотидные последовательности. В ответ на повышение температуры или другие стрессы (например, отравление ядами) вырабатываются особые белки, получившие название *белков теплового шока*. Считается, что быстрое накопление таких белков в клетке обеспечивает физиологическую адаптацию к изменившимся условиям среды. Эти белки чрезвычайно консервативны, они мало менялись в эволюции. Например, белки, имеющие $M_r = 70\,000$ и образующиеся после теплового шока в клетках *E. coli*, растений, насекомых и млекопитающих, проявляют большую степень гомологии по аминокислотной последователь-

ности. Вероятно, они препятствуют образованию осадков денатурированных белков клетки.

Промоторный район гена теплового шока дрозофилы, кодирующего белок с $M_r = 70\,000$, содержит следующую нуклеотидную последовательность:



В ней выделяются районы А и Б. Волнистой чертой отмечена последовательность, необходимая для экспрессии разных генов, кодирующих белки, индуцируемые в условиях теплового шока. Гены, к которым присоединяют этот участок промотора, начинают также активно экспрессироваться при тепловом шоке. В промоторных районах А и Б гена теплового шока дрозофилы подчеркнуты повторяющиеся четырехнуклеотидные мотивы **CTCG** и **GTTC**. Наличие района Б необходимо для полной экспрессии гена. Элементы А и Б, взаимодействующие с белковыми факторами транскрипции, имеют сходные функциональные свойства и обладают синергическим действием, активируя транскрипцию. Гены теплового шока дрозофилы, введенные в клетки млекопитающих, начинают активно экспрессироваться при повышении температуры. Это говорит о том, что не только сами гены теплового шока, но и регуляторные компоненты этой системы генов достаточно консервативны в эволюции.

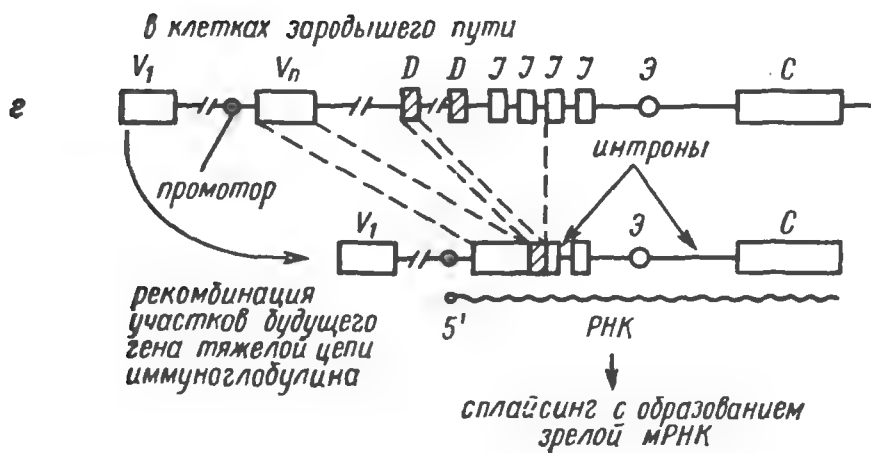
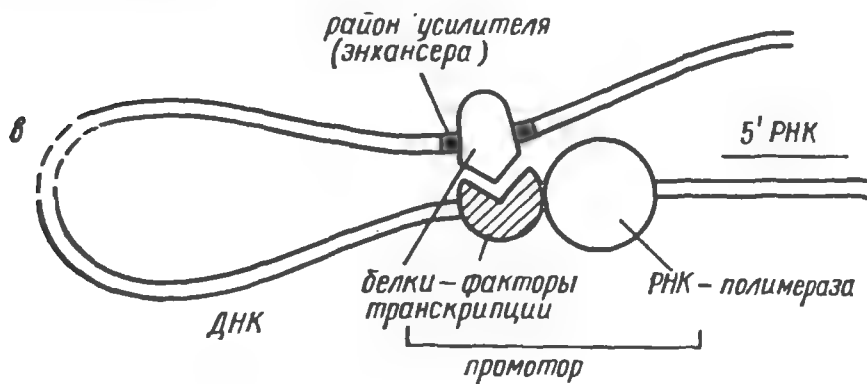
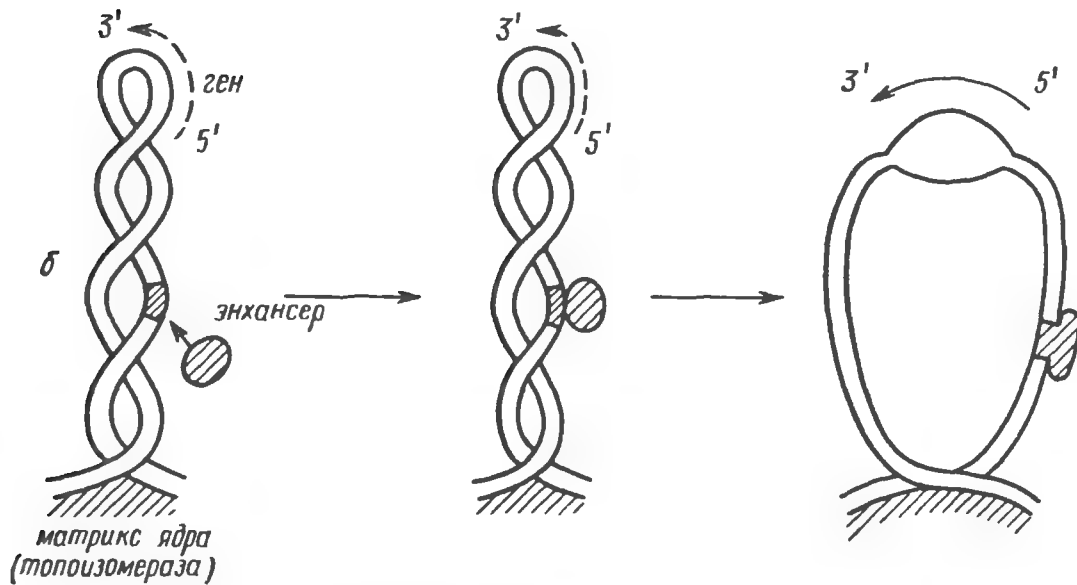
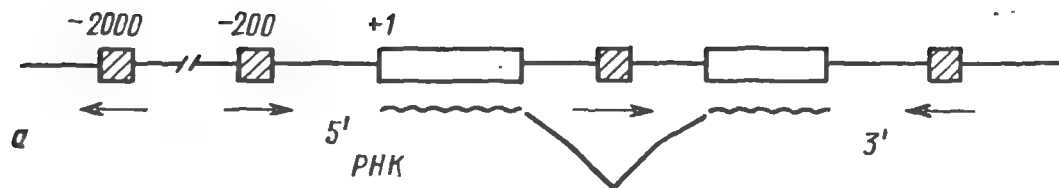
Сложная мозаика цис-действующих регуляторных сигналов в виде перемежающихся коротких «нуклеотидных мотивов» выявлена в промоторной зоне генов, кодирующих металлотионеины млекопитающих (рис. 111, б). Металлотионеин — небольшой, обогащенный цистеином белок, способный связывать двухвалентные металлы. Роль металлотионеина состоит в регуляции концентрации в клетке таких микроэлементов, как цинк и медь, а также в связывании ядовитых тяжелых металлов, например кадмия и ртути. Отравление клеток организма тяжелыми металлами сопровождается накоплением металлотионеина благодаря усилению транскрипции гена (в культурах клеток описаны случаи амплификации этого гена, определяющей их устойчивость к ядам). Геном млекопитающих содержит несколько генов металлотионеина, различающихся особенностями регуляции. В промоторной зоне генов металлотионеинов выявлены повторяющиеся девятинуклеотидные последовательности («мотив **TGCGCTCGG**» или его варианты), наличие которых необходимо для индукции образования металлотионеина в присутствии металлов (рис. 111, б). Если синтетические копии повторяющихся девятинуклеотидных элементов вставить в промоторы других генов, то их экспрессия начнет зависеть от присутствия металлов. Природа непосредственного сигнала (вероятно, белкового комплекса с металлом), ответственного за активацию гена, остается невыясненной. Промоторная зона содержит также «GC-мотивы», по-видимому, обес-

печивающие конститутивный уровень экспрессии гена. Наконец, в составе промотора содержится регуляторный элемент, отвечающий за активацию некоторых металлотионеиновых генов стероидными гормонами — глюкокортикоидами. В принятой схеме действия стероидных гормонов предполагается, что в клетках имеются специфические белки-рецепторы, которые после присоединения к ним гормона-эфектора способны взаимодействовать с геном, индуцируя или усиливая транскрипцию. Были выявлены участки, ответственные за гормонозависимую регуляцию. Последовательность, включающая 15 п. н. и расположенная в положении —250 от *кэп*-сайта (рис. 111, в), отвечает за индукцию глюкокортикоидами гена металлотионеина человека. Таким образом, промоторная область гена металлотионеина представляет собой мозаику регуляторных сигналов в составе ДНК, с которыми взаимодействуют специфические белки. Гипотетическая схема, иллюстрирующая изменение характера взаимодействия белковых факторов в присутствии металлов, представлена на рис. 111, в.

Итак, область эукариотического промотора рассматривается как специфический ДНК-остов, на котором собираются белки транскрипции, узнающие свои сайты связывания и взаимодействующие как друг с другом, так и с РНК-полимеразой. Нельзя исключить, что факторы транскрипции являются ферментами и в процессе этих взаимодействий осуществляются ферментативные модификации как белковых факторов, так и ДНК. Появление нового фактора транскрипции в дифференцированных клетках можно рассматривать как способ включения гена на нужной стадии развития.

Механизм, обеспечивающий разные уровни экспрессии гена в развитии, а также тканеспецифическую экспрессию, может быть основан на использовании разных промоторных участков вблизи гена. Примером может служить ген алкогольдегидрогеназы *D. melanogaster*, обладающий двумя разными промоторными сайтами и двумя последовательностями ТАТА (рис. 111, г). Личиночный промотор работает у поздних эмбрионов и личинок, промотор имаго слабо используется у личинок, но с высокой эффективностью работает у мух. Образующиеся транскрипты подвергаются сплайсингу (рис. 111, г). В результате образуются два типа мРНК, содержащие идентичную нуклеотидную последовательность, кодирующую полипептид, но различающиеся по 5'-лидерным нетранслируемым участкам. Таким образом, регуляция экспрессии гена алкогольдегидрогеназы осуществляется скорее всего не только за счет изменения эффективности образования мРНК с разных промоторов, но и на уровне трансляции, вероятно, благодаря разной ее эффективности, определяемой лидерной зоной мРНК.

В промоторных районах генов дрожжей были выявлены активирующие последовательности (АП), способные усиливать транскрипцию вне зависимости от ориентации по отношению к сайту инициации транскрипции и сохранять это влияние при удалении от него.



Подобные регуляторные элементы, получившие название энхансеров (усилителей), широко распространены в генах многоклеточных эукариот, причем в отличие от генов дрожжей их действие осуществляется не только в положениях перед стартом транскрипции, но и сохраняется при перемещении в 3'-район гена. Оказалось, что ряд из вышеупомянутых нуклеотидных мотивов, обычно обнаруживаемых в промоторном районе, обладают свойствами энхансеров.

2. ЭНХАНСЕРЫ ТРАНСКРИПЦИИ ГЕНОВ, ТРАНСКРИБИРУЕМЫХ РНК-ПОЛИМЕРАЗой II

Опыты с искусственными генными конструкциями, составленными из отрезков ДНК разного происхождения, выявили существование особого цис-действующего элемента регуляции генов эукариот, получившего название усилителя (энхансера) или активатора транскрипции. Энхансеры представлены короткими последовательностями ДНК, состоящими из отдельных элементов (модулей), включающих десятки нуклеотидных пар. Модули могут представлять собой повторяющиеся единицы. Энхансер увеличивает эффективность транскрипции гена в десятки и сотни раз. Впервые энхансеры были обнаружены в составе геномов животных ДНК-содержащих вирусов (SV40 и полиомы), где они обеспечивают активную транскрипцию вирусных генов. Извлеченные из вирусных геномов и включенные в состав искусственных генетических конструкций, они резко усиливали экспрессию ряда клеточных генов. Позднее были обнаружены собственные энхансеры генов эукариотической клетки. Особенность энхансеров состоит в том, что они способны действовать на больших расстояниях (более чем 1000 п. н.) и вне зависимости от ориентации по отношению к направлению транскрипции гена. Оказалось, что энхансеры могут располагаться как на 5'-, так и на 3'-конце фрагмента ДНК, включающего ген, а также в составе интронов (рис. 112, а). Например, энхансеры были выявлены в районе 400 п. н. перед стартом транскрипции генов инсулина и химо-трипсина крысы. В случае гена алкогольдегидрогеназы дрозофилы энхансер был локализован за 2000 п. н. перед промотором. Энхансеры обнаружены на 3'-фланге гена, кодирующего полипептидный гормон-плацентарный лактоген человека, а также в составе интронов генов иммуноглобулинов и коллагена.

Рис. 112. Энхансеры эукариотических генов:

а — возможные районы локализации энхансера (заштрихованный прямоугольник), с разной ориентацией (стрелки) относительно направления транскрипции; б — взаимодействие белка с энхансером; стрелка — транскрипция гена; в — сгибание молекулы ДНК, обеспечивающие контакт усилителя с промотором при участии белковых факторов транскрипции; г — усилитель в функционирующем гене тяжелой цепи иммуноглобулина лимфоидных клеток; прямоугольники — экзоны; д — энхансер, локализованный в большом интроне

Энхансеры могут содержать разные нуклеотидные последовательности, составленные из нескольких «нуклеотидных мотивов», каждый из которых обладает указанными особыми свойствами. Такие «мотивы» (модули) могут быть повторены в одном сайте или чередоваться друг с другом. В составе определенного семейства энхансеров (например, в наиболее изученных энхансерах геномов вирусов) можно выделить отдельные общие «мотивы», нуклеотидные замены в которых приводят к резкому снижению их биологической активности. Например, резко падает активность энхансера с последовательностью $TG\overset{*}{G}AAAC$ в результате замены гуанилового нуклеотида, отмеченного звездочкой.

Принципы действия энхансеров, способных оказывать свое влияние на значительном расстоянии (более чем тысячи нуклеотидных пар) и вне зависимости от ориентации по отношению к старту транскрипции, не выяснены. Короткие нуклеотидные блоки могут служить центрами связывания специфических ядерных белков, выступающих как транс-действующие факторы. Сила энхансера, вероятно, может зависеть от числа таких блоков (модулей). Обсуждаются следующие два основных механизма действия энхансеров. Считается, что функциональные участки генома, содержащие один или несколько генов, образуют длинные петли, включающие десятки тысяч нуклеотидных пар ДНК. Высказано предположение, что петли закреплены в матриксе клеточного ядра и могут быть сверхспирализованы. В состав матрикса входит топоизомераза II, по-видимому, определяющая топологию петли ДНК (см. гл. XII). В таком случае взаимодействие энхансера с белками может менять конформацию всей петли, включая и удаленный от энхансера участок ДНК, в результате чего в составе петли изменяется локальная структура хроматина и облегчается транскрипция гена (рис. 112, б). Более вероятно, что влияние энхансера, связанного с белком, определяется его непосредственным взаимодействием с РНК-полимеразой и другими факторами транскрипции в процессе инициации. Такое взаимодействие может осуществляться благодаря сгибанию молекулы ДНК, что создает возможность непосредственного контакта районов промотора и удаленного от него энхансера, связанных со специфическими белками (рис. 112, в).

Способность ряда энхансеров взаимодействовать со специфическими белками дифференцированной клетки, вероятно, обеспечивает их важное свойство — тканевую специфичность. Тканеспецифический энхансер впервые был выявлен в генах, кодирующих тяжелую полипептидную цепь иммуноглобулинов. При образовании функционирующего гена иммуноглобулина происходит программированная в развитии рекомбинация генетического материала. Один из нескольких сотен геномных сегментов, кодирующих варьирующую часть молекулы антитела (V-гены), в результате последовательных рекомбинационных процессов соединяется с D- и J-эlemen-

тами и, наконец, с районом, кодирующим константные С-районы иммуноглобулина и содержащим энхансер (рис. 112, а). В процессе транслокаций каждый V-ген сохраняет свой промотор. Однако только промотор, перенесенный к энхансеру и находящийся в составе функционально перестроенного гена, активируется в В-лимфоцитах, где ген должен работать. Энхансер, свойства и функции которого были доказаны с помощью искусственных генетических конструкций, располагается в районе интрона. Заметим, что приближение усилителя к V-генам — это лишь один из многих этапов в сложном и многоступенчатом процессе активации экспрессии генов иммуноглобулинов. Энхансер обнаруживает выраженную тканевую специфичность: его активность проявляется в лимфоидных клетках, но не в фибробластах или клетках легочной ткани. Энхансеры с тканеспецифическим действием описаны и для других генов, например для генов инсулина и химотрипсина, которые не подвергаются перестройке в процессе клеточной дифференцировки.

Изменение характера экспрессии генов можно наблюдать в бесклеточных системах, компонентами которых являются собственно ген, энхансер и специфические регуляторные белки. Оказывается, что энхансер в зависимости от добавляемого белкового фактора может начать вести себя и как негативно действующий «глушитель» (англ. *silencer*) экспрессии гена. Негативное действие такого элемента, проявляется при связывании с тканеспецифичным трансдействующим белковым фактором. Подобные элементы с негативными эффектами были обнаружены, например, вблизи генов инсулина I и α -фетопротеина крысы, экспрессия которых наблюдается соответственно в β -клетках поджелудочной железы и в печени, но отсутствует во многих других тканях. Негативное действие «глушителей», как и в случае энхансеров, не зависит от положения и ориентации относительно сайта инициации транскрипции.

По мере исследования регуляторных элементов генома эукариот оказалось, что не удастся провести строгих функциональных различий между энхансерами и вышеописанными элементами промотора. Прежде всего ориентация таких активных элементов промотора, как «GC-мотивы», может быть любой относительно направления транскрипции гена. Более того, промоторные элементы генов металлотioneина и генов теплового шока, отделенные от ТАТА-последовательности, определяющей сайт инициации транскрипции, также сохраняют специфические регуляторные свойства вне зависимости от ориентации. Промоторный участок гена металлотioneина, включающий ряд регуляторных мотивов, способен сильно стимулировать экспрессию гена β -глобина и в том случае, если промотор расположить на 3'-фланге относительно направления транскрипции, причем на значительном расстоянии от β -глобина. Примечательно, что экспрессия β -глобина теперь будет зависеть от ионов металлов. В результате регуляторный участок гена металлотioneина можно рассматривать не только как промотор, но и как *индуцируемый*

энхансер активности гена. Подобным образом специфический промотор генов теплового шока, отделенный от мотивов ТАТА и ССААТ и находящийся на значительном расстоянии от старта транскрипции, вне зависимости от ориентации по отношению к нему активирует ген β -глобина при повышении температуры.

Таким образом, регуляторные элементы генов, которые первоначально относили либо к промоторным, либо к энхансерам, обладают рядом общих функциональных характеристик. На физической карте регуляторной области гена они могут располагаться в одном районе, создавая сложную мозаику регуляторных сигналов. Примеры с генами металлотионеина или белков теплового шока показали, что цис-действующие регуляторные элементы способны активно выполнять свою роль при рекомбинации с гетерологичными или с гомологичными генами организмов, отстоящих в эволюционном плане достаточно далеко друг от друга. Следовательно, взаимодействие этих элементов и соответствующих белковых транс-действующих регуляторов генной активности основано на достаточно общих консервативных принципах, сохраняющихся в процессе эволюции. Возможность длительной эволюционной сохранности цис- и транс-действующих компонентов, отвечающих за строгую тканевую и временную экспрессию генов, а следовательно, за регуляцию развития организма, неоднократно была показана экспериментально. Так, например, демонстративны опыты с трансгенными особями *D. melanogaster*, геном которых содержал гены шелкопряда, которые кодируют комплекс белков оболочки яйца и работают в фолликулярных клетках, окружающих яйцо. В фолликулярных клетках трансгенных мух гены шелкопряда экспрессировались в то же самое время развития, что и собственные гены. Считается, что общий предок дрозофилы и шелкопряда существовал 240 млн. лет назад — в то время, когда только началась дивергенция птиц и млекопитающих. Возможно, эволюция в значительной степени шла не путем создания принципиально новых типов взаимодействия регуляторных элементов, а путем их перекомбинации и возникновения новых связей между генами. Отсутствие требований к строгой локализации и ориентации регуляторных последовательностей относительно генов сильно расширяет возможности создания новых регуляторных связей путем перетасовки регуляторных элементов в процессе эволюции. В таком случае эволюция регуляторных элементов эукариот может быть основана на изменчивости мозаики ограниченного числа регуляторных нуклеотидных мотивов.

Чрезвычайно высокая степень консервативности во взаимодействиях белков транскрипции с гетерологичными промоторами и энхансерами, а также белков транскрипции разного происхождения друг с другом была показана следующими экспериментами. Добивались экспрессии клонированного гена дрожжей в клетках млекопитающих и следили за функцией продукта этого гена — белка GAL4 (рис. 111, а). Оказалось, что белок GAL4, образующийся в клетках

млекопитающих, связывается с родственной энхансерной последовательностью в ядре и активирует транскрипцию генов млекопитающих. Другими словами, дрожжевой белок GAL4 способен не только узнавать энхансерные последовательности, сходные с АП-последовательностями генов дрожжей, но и достаточно эффективно взаимодействовать с факторами транскрипции и РНК-полимеразой млекопитающих. Возможность этих взаимодействий объясняется структурным сходством по крайней мере отдельных участков активаторных белков дрожжей и белков млекопитающих, связывающихся с энхансерами и другими белками транскрипции, а также высокой консервативностью аминокислотной последовательности большой субъединицы РНК полимеразы II у всех эукариот. Следовательно, способность к функциональным контактам гетерологических белков, участвующих в транскрипции, сохраняется в эволюции эукариотических организмов.

3. ТРАНСКРИПЦИЯ РИБОСОМНЫХ ГЕНОВ С ПОМОЩЬЮ РНК-ПОЛИМЕРАЗЫ I

Более половины валового синтеза РНК эукариотической клетки приходится на образование 18S, 28S и 5,8S РНК рибосом. Рибосомальные гены представлены сотнями повторяющихся копий, сгруппированных в одном или нескольких участках генома. Многократная повторяемость генов рРНК, а также их амплификация на определенных стадиях развития отражают потребность клетки в больших количествах транскриптов этих генов. Блок активно работающих рибосомных генов формирует особую структуру — ядрышко, где осуществляется синтез РНК. Гены, кодирующие рибосомные РНК, разделены промежутками — спейсерами (англ. spacer), длина и структура которых может достаточно сильно различаться даже у близких видов (рис. 113). Ранее считали, что районы спейсера не транскрибируются. Действительно, электронно-микроскопические картины транскрипции рибосомных генов, казалось бы, подтверждали этот вывод. Однако оказалось, что транскрипты спейсеров подвергаются особенно быстрому процессингу и деградации. В конце спейсера находится сайт терминации Т, включающий консервативную последовательность из нескольких нуклеотидов. Функциональная особенность этого сайта заключается в том, что он же отвечает и за реинициацию нового транскрипта. Одно и то же мутационное изменение в этом районе спейсера снижает как эффективность терминации, так и реинициации транскрипции. При освобождении транскрипта в сайте Т полимераза сохраняется связанной с ДНК. Взаимодействуя с дополнительным белковым фактором, полимераза тут же настраивается на следующий акт реинициации транскрипции. Благодаря сопряжению актов терминации и инициации в одном сайте обеспечивается высокая эффективность процесса транскрипции, поскольку каждый новый акт инициации не требует поиска промотора.

В спейсере обнаружены повторяющиеся последовательности, которые резко, в десятки раз, активируют транскрипцию. Подобно энхансерам генов, транскрибируемых РНК-полимеразой II, такие повторы, включающие по 40—60 п. н., активируют транскрипцию вне зависимости от их ориентации по отношению к направлению транскрипции. Эффекты таких повторов также мало зависят от расстояния между ними и стартом транскрипции. Интенсивность транскрипции пропорциональна числу таких усилителей в составе спейсера. Однако механизмы, с помощью которых энхансеры, локализованные в районе транскрибируемого спейсера, оказывают влияние на транскрипцию рибосомных генов, могут быть совсем иными по сравнению с механизмом действия энхансеров генов, транскрибируемых РНК-полимеразой II. Повторяющиеся по длине спейсера элементы способствуют связыванию молекул РНК-полимеразы I с ДНК-матрицей в районе спейсера. Предполагается, что спейсер наряду с молекулами РНК-полимеразы, осуществляющими транскрипцию, может содержать связанный пакет молекул полимеразы, тесно расположенных по длине спейсера и продвигающихся к промотору. В районе промотора осуществляется инициация транскрипции этими молекулами РНК-полимеразы. В результате тандемная организация рибосомных генов, разделенных длинными спейсерами, обеспечивающими связывание полимеразы и содержащими регуляторные сайты транскрипции, предназначена для обеспечения высокой эффективности функционирования рибосомных генов. Эффективность процесса, по-видимому, также значительно увеличивается

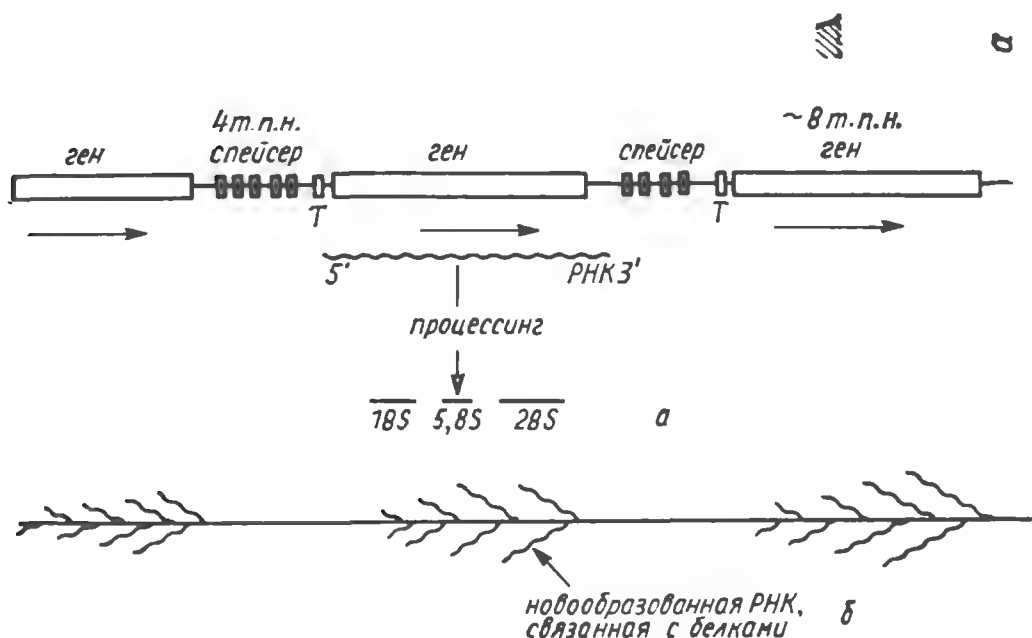


Рис. 113. Транскрипция рибосомных генов эукариот РНК-полимеразой I (а) и картины транскрипции, выявляемой методами электронной микроскопии (б)

Черные прямоугольники — положение элементов, активирующих транскрипцию; Т — сайты терминации (реинициации)

благодаря компартиментализации (концентрированию) всех компонентов системы в одной структуре — ядрышке.

Различия в структуре спейсеров у близких видов очень велики, поэтому спейсеры рассматривались как функционально нейтральные. Обнаружение важных регуляторных элементов в составе спейсера позволило поставить под сомнение общее представление о том, что в процессе эволюции быстро меняются те районы ДНК, которые функционально нейтральны. Напротив, можно предполагать, что здесь осуществлялась сравнительно быстрая совместная эволюция («коэволюция») функционально важных участков ДНК спейсера и белков, взаимодействующих с ними.

Транс-действующие факторы транскрипции, связывающиеся с элементами промотора РНК-полимеразы I, не изучены. Значительно больше известно о структуре и механизмах действия белковых факторов транскрипции, взаимодействующих с РНК-полимеразой III.

4. РЕГУЛЯЦИЯ ТРАНСКРИПЦИИ, ОСУЩЕСТВЛЯЕМОЙ РНК-ПОЛИМЕРАЗой III

РНК-полимераза III транскрибирует гены 5S РНК, тРНК и 7S РНК. Гены 5S РНК, как и гены других рибосомальных РНК, обычно представлены сотнями или даже тысячами tandemно повторяющихся копий, разделенных спейсерами (рис. 114). Эти гены собраны в одном или нескольких районах генома. Гены разных тРНК (в эукариотической клетке насчитывается 40—60 основных типов тРНК) также часто сгруппированы в кластеры, расположенные в разных хромосомах. Характер организации генов тРНК в составе

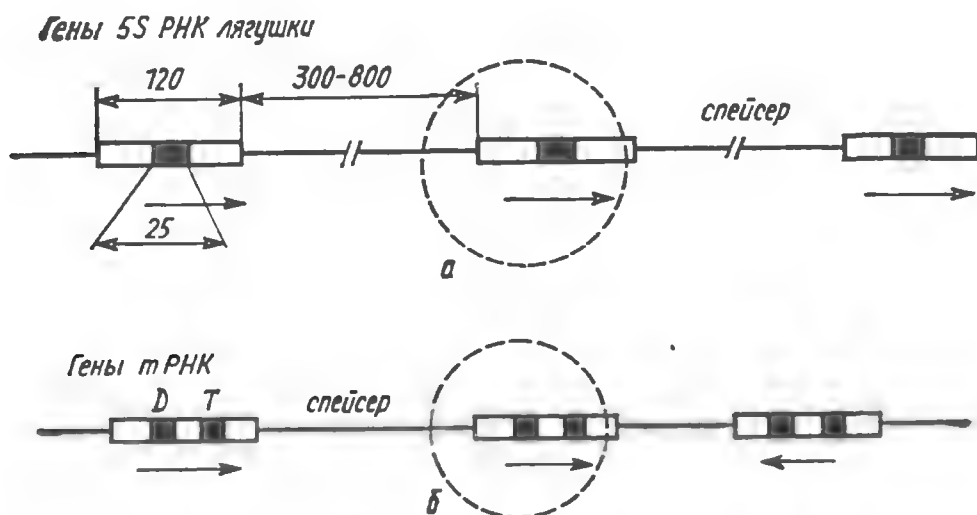


Рис. 114. Гены 5S РНК (а) и тРНК (б), транскрибируемые РНК-полимеразой III:

Пунктир — РНК-полимераза; стрелки — направление транскрипции; зачернены внутренние участки генов, контролирующих транскрипцию (размеры указаны в нуклеотидах)

таких блоков сильно варьирует у разных организмов. В целом организация генов тРНК в кластерах не столь регулярна, как в случае рибосомальных РНК: направление транскрипции близлежащих генов бывает противоположным, а гены могут быть разделены нерегулярными промежутками, включающими 100—150, а иногда и несколько тысяч нуклеотидных пар. Природа и функции участков ДНК, разделяющих гены тРНК, не выяснены.

Особенность участков генов 5S РНК и тРНК, узнаваемых факторами транскрипции и РНК-полимеразой и обеспечивающих инициацию транскрипции, состоит в том, что они локализованы непосредственно в транскрибируемом районе (или районах) гена. Так, район внутри генов 5S РНК лягушки, кодирующий нуклеотиды с 55-го по 80-й, необходим для инициации транскрипции (рис. 114). Размеры транскрипционного комплекса, включающего субъединицы полимеразы, достаточно велики, чтобы взаимодействовать с внутренними районами небольшого по размерам гена и одновременно инициировать транскрипцию. Сигналом терминирования транскрипции служит последовательность из нескольких следующих друг за другом тимидиловых нуклеотидов. Роль спейсера в транскрипции, по крайней мере в опытах *in vitro*, пока не была обнаружена. Новообразованная 5S РНК, по-видимому, не подвергается модификации и процессингу.

Внутри генов тРНК обнаружены два района, необходимые для того, чтобы осуществилась правильная и эффективная транскрипция. Например, для одного из исследованных генов тРНК это области, кодирующие нуклеотиды 8—25 и 50—58 молекулы тРНК. Два участка такого «расщепленного промотора» совпадают с районами гена, которые соответствуют наиболее консервативным районам молекул тРНК в дигидроуридиловой (D) и псевдоуридиловой (T) петлях, соодержащих наибольшее количество инвариантных нуклеотидов. Для активной транскрипции требуется сохранность расстояния между районами D и T промотора, которое составляет около 50 п. н. (рис. 114). В то же время эту область гена можно заменить без ущерба на чужеродную. Район генов тРНК, кодирующий интрон, выщепляемый в результате процессинга (см. гл. VIII), не играет роли в инициации транскрипции. Внутри районов генов 5S РНК и тРНК, определяющих точную и активную транскрипцию, можно найти гомологии по нуклеотидной последовательности.

Транскрипция генов 5S РНК и тРНК осуществляется с участием выделенных и очищенных белков — факторов транскрипции. Особенно хорошо изучен специфический фактор транскрипции TF III A (англ. transcription factor) 5S-генов. Фактор представляет собой полипептид с $M_r=40\,000$, он связывается с внутренним контролирующим элементом 5S-гена. Вслед за ним связываются два других белка и присоединяется РНК-полимераза. Одна из особенностей белка TF III A состоит в том, что он специфически связывается не только с ДНК, но и с 5S РНК. Поэтому при большой кон-

центрации продукта транскрипции — 5S РНК — фактор связывается, а экспрессия 5S-генов подавляется по принципу обратной связи.

Белок TF III A был первым эукариотическим регуляторным полипептидом транскрипции с известной аминокислотной последовательностью, для которого удалось построить доменную структурную модель. В этом белке выявлены 9 повторяющихся, но отличающихся друг от друга доменов — «пальцев», каждый из которых включает около 30 аминокислот. Домены содержат инвариантные участки, включающие два цистеиновых и два гистидиновых остатка, связанных с ионом цинка (рис. 115). «Концы» разных «пальцев» (петли) несут варьирующие аминокислотные остатки, среди которых встречаются положительно заряженные, которые, по-видимому, способны легко взаимодействовать с ДНК. Как оказалось, подобная структура регуляторного белка закодирована в ряде других генов, кодирующих регуляторные белки эукариот. Так, ген *Krüppel* (калека), контролирующий развитие дрозофилы, кодирует белок, содержащий четыре подобных домена. Такие домены обнаружены и в белках — рецепторах гормонов. Предполагается, что выступающие связывающиеся с ДНК разные пальцы, соединенные друг с другом гибкими мостиками, осуществляют сразу несколько контактов с ДНК. Такая модель строения TF III A позволяет предполо-

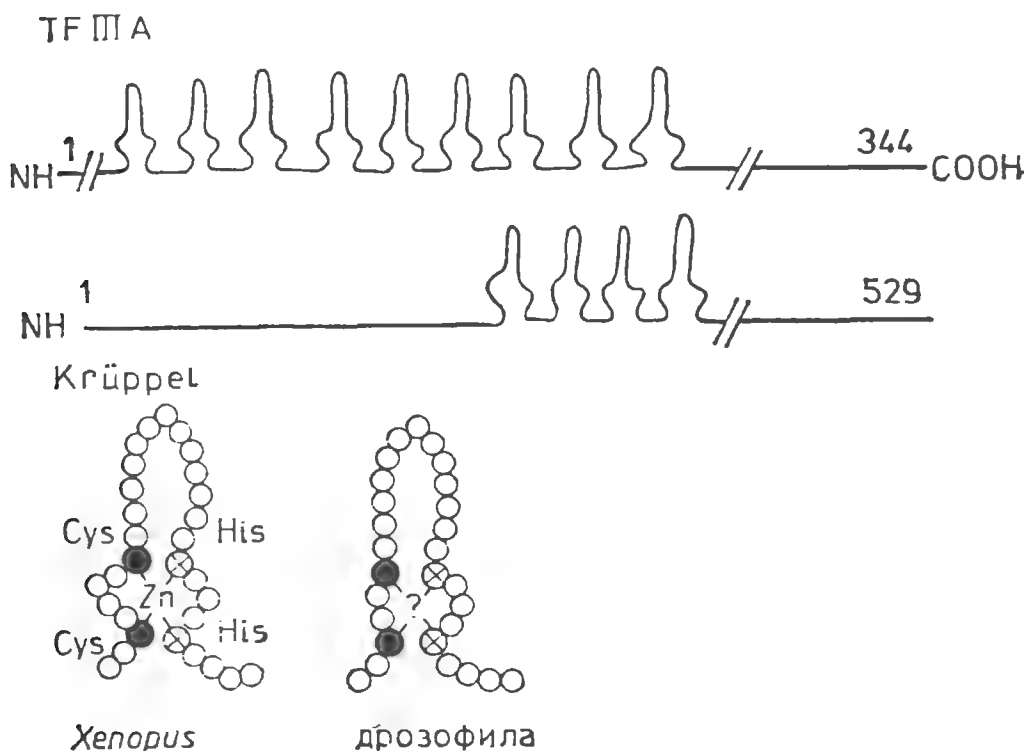


Рис. 115. Схема линейного расположения доменов-«пальцев» в составе фактора транскрипции TFIIIA и белка, кодируемого геном *Krüppel*, управляющим развитием дрозофилы

Цифры — номера аминокислотных остатков; зачерненные и обозначенные крестиками кружки — соответственно расположение остатков цистеина и гистидина

жить, почему фактор сохраняется связанным с ДНК при повторных актах транскрипции гена: при продвижении полимеразы одни пальцы сохраняют контакты, тогда как другие, переставая взаимодействовать с ДНК, обеспечивают продвижение РНК-полимеразы. «Пальцы», различающиеся по силе связывания с ДНК или РНК, по-видимому, обеспечивают высокоспецифическую систему узнавания нуклеиновых кислот белком.

Были рассмотрены три группы эукариотических генов, транскрипция которых осуществляется разными РНК-полимеразами при участии белковых факторов, взаимодействующих с характерными для каждой группы регуляторными элементами. Однако кроме них существуют еще «гибридные» системы транскрипции, в которых, по-видимому, одновременно могут использоваться способы регуляции, представленные в каждом из рассмотренных типов транскрипции. Так, РНК-полимераза III транскрибирует гены малых ядерных РНК (см. гл. VIII) типа U6, а также гены 7SK РНК неизвестной функции, хотя те и другие не содержат «внутренних» промоторов и, напротив, на 5'-конце несут ряд элементов, характерных для систем транскрипции с помощью РНК-полимеразы II.

5. ГЕНЫ, НАПРАВЛЯЮЩИЕ РАЗВИТИЕ ОРГАНИЗМОВ

Регуляция экспрессии генов эукариот лежит в основе программы развития многоклеточных организмов. В начале XX в. благодаря работам Т. Моргана и Э. Вильсона стало очевидным, что развитие программируется генами. Поэтому один из подходов к пониманию закономерностей высокоупорядоченного процесса развития состоит в выявлении генов, контролирующих ключевые стадии развития.

Закономерности, управляющие развитием организмов, которое сопровождается сложной последовательной дифференцировкой тканей и формообразованием, остаются загадкой. Один из путей их исследования основывается на молекулярном анализе структуры и экспрессии генов, отвечающих за критические этапы развития дифференцировки и морфогенеза. Лишь немногие организмы, как, например, *D. melanogaster*, пригодны для исследования развития методами молекулярной генетики, поскольку гены, контролирующие развитие, можно выявить лишь тогда, когда хорошо разработана генетика объекта исследования.

Этапы развития дрозофилы подробно описаны. После оплодотворения и слияния материнского и отцовского ядер у дрозофилы происходят последовательные синхронные деления ядра зиготы (рис. 116, 1—6). Образуется многоядерный, не имеющий клеточного строения синцитий, содержащий несколько сотен ядер. Ядра синцития эквивалентны по своим потенциям в развитии (тотипотентны), т. е. каждое ядро еще сохраняет способность быть предшественником любой ткани организма. Эти ядра еще не детерминированы.

Затем ядра мигрируют к поверхности эмбриона в зону кортикальной плазмы, где делятся и образуют монослой, в состав которого входит около 6000 ядер. Это стадия бластодермы, на которой начинается образование клеточных мембран и осуществляется детерми-

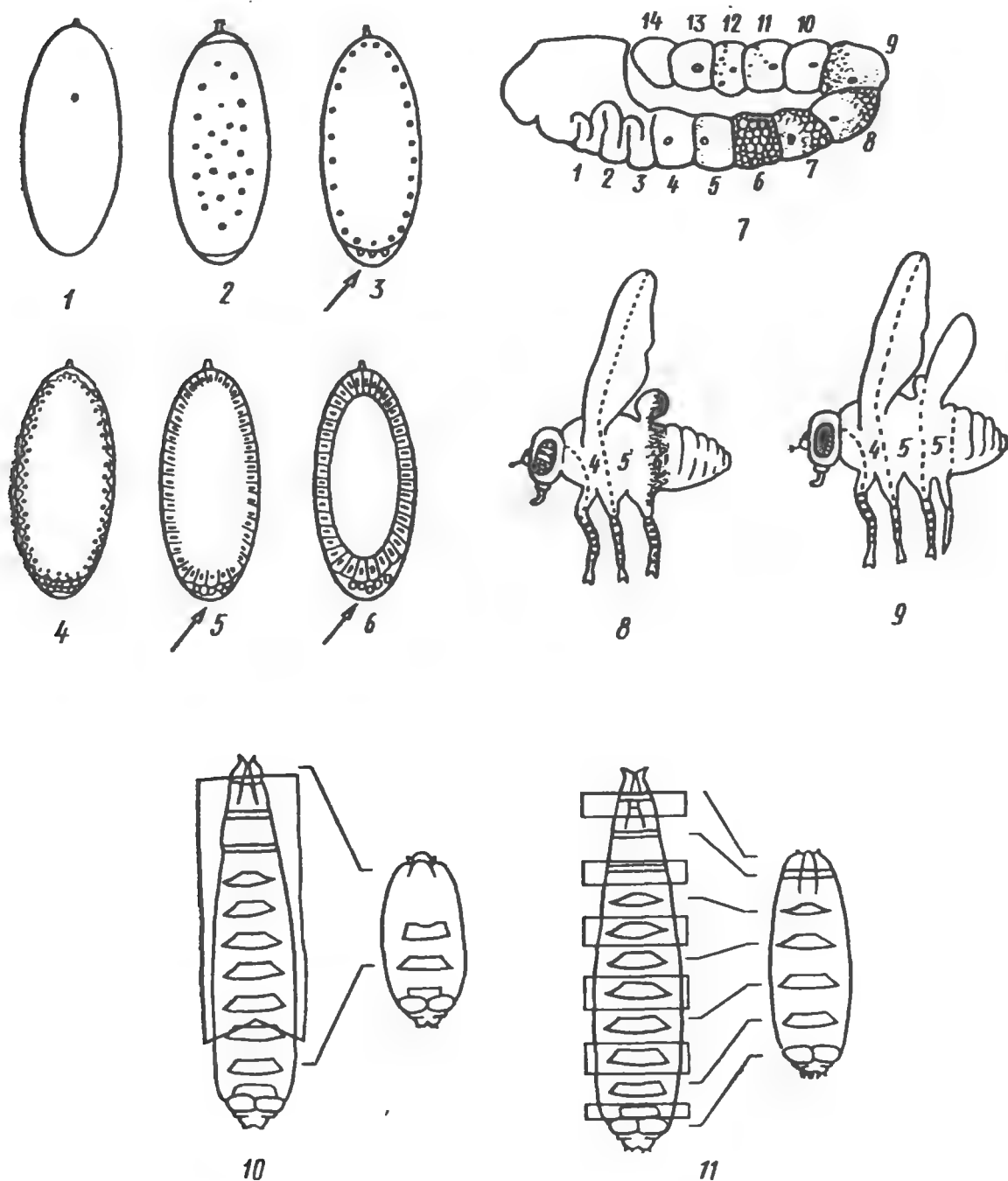


Рис. 116. Этапы раннего развития дрозофилы и его нарушения, вызываемые мутациями:

1 — зигота; 2 — деление ядер и образование синцития; 3 — миграция ядер; 4, 5 — образование клеточных структур (целлюляризация); 6 — бластодерма (стрелки указывают положение обособляющихся клоток зародышевого пути); 7 — сегментация эмбриона и личинки (в сегменте 6 накапливаются транскрипты гена *bx*); 8 — сегменты тела нормальной мухи; 9 — превращение (трансдетерминация) сегмента 6 в сегмент 5 в результате гомеотических мутаций в локусе *BX*; 10, 11 — нарушения сегментации, наблюдаемые у личинок (отсутствующие сегменты тела обведены): 10 — одновременная потеря группы прилегающих друг к другу сегментов; 11 — потеря «четных» сегментов

нация ядер и клеток. Судьба клеток уже строго определена в зависимости от их положения в кортикальной цитоплазме — эти клетки предназначены для образования тех или иных сегментов тела личинки или взрослого насекомого.

Самые ранние стадии развития дрозофилы, когда устанавливаются так называемые пространственные координаты эмбрионов, определяющие передний и задний или брюшной и спинной отделы, контролируются группой генов матери. Эти гены функционируют на стадии образования яйца, и их продукты неравномерно распределяются по яйцеклетке. Предполагается, что материнские гены и их продукты обеспечивают *позиционную информацию*, которая воспринимается генами, работающими после оплодотворения, в зиготе. Представление о наличии в цитоплазме яйца позиционной информации, определяющей направление развития групп эмбриональных клеток, подчеркивает роль взаимного влияния частей будущего эмбриона в развитии, но никак не вскрывает природы этих взаимодействий. Мутации в генах, определяющих структуру неоплодотворенного яйца, оказывают так называемый материнский эффект, нарушая развитие эмбриона. Например, структуры, свойственные данному району, заменяются иными, характерными для других районов развивающегося организма. Вероятно, такие материнские гены оказывают свое действие на стадии ядерного синцития, до образования клеток бластодермы, когда диффузия продуктов генов затрудняется в результате образования клеточной мембраны. Транскрипты таких генов локализуются в соответствующих отделах (например, переднем или заднем) неоплодотворенного яйца или развивающегося эмбриона.

Вторая группа генов отвечает за последующие морфологические изменения структуры эмбриона — формирование сегментов тела личинки, которые прослеживаются и у имаго (рис. 116, 7—9). Мутации в одних генах приводят к выпадению нескольких соседних сегментов (рис. 116, 10), в других — к выпадению участков каждого сегмента (рис. 116, 11). Например, мутация *fushi tarazu (ftz)*, что по-японски означает недостаточное число сегментов, сопровождается утратой передней половины каждого нечетного сегмента и задней половины каждого четного. Оставшиеся половины сегментов сливаются, и в результате мутантный эмбрион несет лишь половину нормального числа сегментов. Все такие мутации в гомозиготном состоянии приводят к гибели эмбриона.

Следующий важный этап развития заключается в установлении морфологических различий между сегментами, которые становятся зачатками головы, брюшка и грудных сегментов. Эти процессы контролируются третьей группой генов дрозофилы. Мутации в этих генах приводят к ошибкам при формообразовании сегментов тела: например, головной сегмент может образовать ногу, а один из грудных, с которым в норме связано жужжальце, вторую пару крыльев. В результате нормально (или почти нормально) сформиро-

ванная структура возникает в неподходящем месте (рис. 116, 8, 9). Такие мутации, нарушающие нормальные пути развития сегментов и вызывающие фактически превращение одного сегмента в другой, получили название *гомейотических* или *гомейозисных* (homoeotic, или homeotic) от термина «homoeosis», предложенного английским генетиком Бетсоном для обозначения случаев подобных морфологических трансформаций в развитии. В результате комбинаций разных аллелей гомейотических мутаций в локусе ВХ удастся сконструировать генотипы, определяющие, например, развитие четырехкрылых мух с четырьмя парами ног (рис. 116, 9).

Ряд гомейотических генов и генов сегментации был успешно клонирован благодаря тем ухищрениям, которые, с одной стороны, предоставила технология получения рекомбинантных ДНК, а с другой — классическая генетика, накопившая бесценный багаж знаний о точной локализации этих генов в хромосомах. Во-первых, исследование структуры клонированных генов развития показало, что они кодируют белки, связывающиеся с ДНК. Во-вторых, наличие клонированных генов позволило изучить характер пространственной экспрессии этих генов. Если гены, играющие важную роль в процессе развития, клонированы, то можно выявить те клетки и районы тела эмбриона, в которых они экспрессируются. Для исследования транскрипции гена срезы тканей эмбриона гибридизуют с мечеными фрагментами ДНК клонированного гена. В тех клетках, которые содержат образующиеся в ядре и накапливающиеся в цитоплазме транскрипты гена, будут образовываться гибриды с участием меченой ДНК, комплементарной этим транскриптам. С помощью автордиографии после проведения такой гибридизации *in situ* можно выявить клетки, в которых исследуемый ген интенсивно транскрибируется (рис. 116, 7 и 117, а). Клетки, не содержащие этих транскриптов, остаются практически не мечеными. Оказалось, что транскрипция генов сегментации выявляется в клетках еще не сегментированного эмбриона в виде перемежающихся полос шириной в несколько клеток. После радиоавтографии эмбрион раскрашивается «под зебру». Например, в нормальных эмбрионах обнаруживается семь отдельных полос гибридизации с геном *ftz*. Гибридизация наблюдается как раз с теми участками сегментов, которые отсутствуют у мутантных личинок. Вероятно, нарушение транскрипции гена *ftz* приводит к гибели клеток — предшественников отдельных частей сегментов. Пространственное распределение транскриптов ограничено теми районами, в которых функционирование соответствующих генов, вероятно, требуется для нормального развития.

Таким образом, на стадии бластодермы, когда еще не наблюдается никаких морфологических намеков на сегментацию эмбриона, удастся выявить работу генов, приуроченную к определенным чередующимся зонам бластодермы, детерминированным на образование определенных сегментов тела. Этот результат эксперименталь-

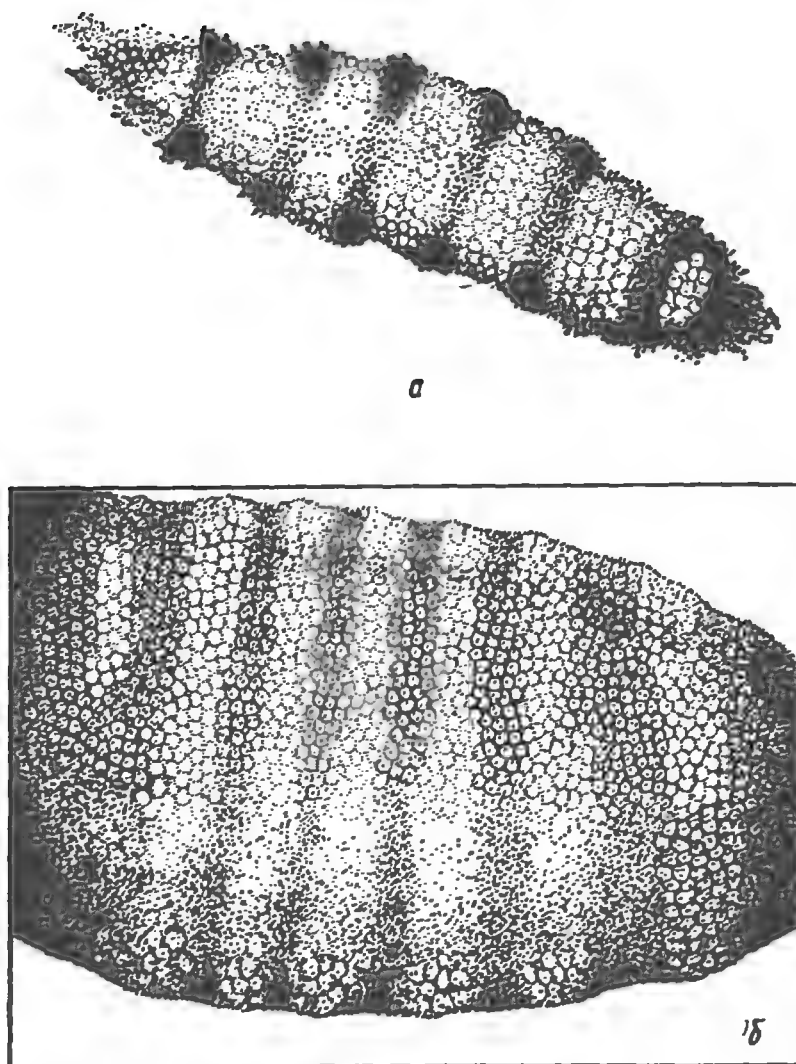


Рис. 117. Экспрессия гена *ftz*, определяющего сегментацию у *D. melanogaster*, на ранних стадиях развития в эпителии бластодермы:

a — распределение транскриптов (гибридизация с меченой ДНК); *б* — распределения белка, выявляемого методом иммунофлуоресценции

но подтверждает представление о том, что состояние детерминации отражает установление определенного «профиля» генной активности, что предшествует появлению тканевых морфологических различий. Подобным образом экспрессию гомеотического гена определяющего морфогенез грудного и брюшного сегментов имаго, удастся обнаружить только в тех морфологически недифференцированных сегментах эмбриона, которые являются предшественниками соответствующих частей тела имаго (см. рис. 116, 7).

Пространственная экспрессия генов сегментации строго регулируется в раннем эмбриогенезе. Ответ на вопрос о том, каким образом достигается пространственная экспрессия этих генов, тесно связан с решением проблемы закономерностей развития.

Белковые продукты ряда генов развития были локализованы в ядрах эмбриональных клеток. Метод локализации состоит в следую-

щем. Если известна нуклеотидная последовательность клонированного гена, то в ее составе легко найти районы «открытых рамок трансляции», предназначенных для кодирования полипептида. Подставляя кодирующие области гена (экзоны) под сильный прокариотический промотор, в клетках бактерий удастся наработать соответствующий белок в заметных количествах. Удобно воспользоваться искусственным гибридным геном, который активно экспрессируется в бактериальных клетках и содержит не только экзоны исследуемого гена, но и, например, участок гена бактериальной β -галактозидазы, придающей химерному белку ферментативную активность. В таком случае за очисткой гибридного белка можно следить по ферментативной активности. Очищенный антиген используют для получения антител против той части молекулы, которая кодируется эукариотическими экзонами. В результате с помощью флуоресцирующих антител удастся выявить в ядре или цитоплазме клеток белки с неизвестной функцией, но специфическими антигенными детерминантами, кодируемыми «генами развития» (рис. 117, б).

Гены развития дрозофилы, причисленные к трем отдельным группам, сложным образом взаимодействуют между собой. Так, гены матери, определяющие пространственные координаты яйца, влияют на характер экспрессии генов сегментации, нарушая образование характерного рисунка «под зебру», контролируемого геном *ftz*.

Гены, контролирующие развитие дрозофилы, могут быть достаточно протяженными и включать многие десятки тысяч нуклеотидных пар ДНК. Например, длина первичного транскрипта локуса *VX* очень велика, она достигает 70 т. п. н. В состав ряда таких генов входят необычно длинные интроны. Возможно, альтернативные пути сплайсинга (процессинга) длинных первичных транскриптов локуса *VX* приводят к образованию целого набора мРНК, которые транслируются в разных группах клеток с образованием белков, несущих разные функции. Однако детали строения этих генов и особенности регуляции их активности еще мало исследованы.

При изучении молекулярной структуры как гомеотических генов, так и генов сегментации в составе некоторых экзонов был выявлен общий гомологичный фрагмент, включающий около 180 п. н. Этот фрагмент, получивший название *гомеоблока* (англ. homeobox), составляет 10—20 % от всей последовательности гена, кодирующей белок. Вероятно, действие отбора было направлено прежде всего на поддержание консервативной структуры «гомеодомена» белка, кодируемого гомеоблоком. Домен обогащен аргинином и лизином. Его полипептидная структура, судя по расположению аминокислот, сходна с той, которая обнаруживается в регуляторных ДНК-связывающихся белках прокариот и дрожжей (например, белка *Сго фага* λ), где она представлена двумя согнутыми под углом α -спиралями. Скорее всего, гомеоблок определяет специфические взаимодействия белка с регуляторными сайтами ДНК. Возможно, «гомеотические

белки» регулируют специфический класс генов, содержащих одинаковые или сходные последовательности ДНК.

Высокая консервативность гомеоблока проявляется не только при исследовании разных генов развития у дрозофилы. Перекрестная гибридизация с геномами червей, иглокожих и позвоночных, включая человека, выявила наличие нескольких фрагментов геномной ДНК, гибридизующихся с гомеоблоком дрозофилы. Если гены этих организмов, содержащие подобные последовательности, также вовлечены в регуляцию ключевых этапов развития, то гомеоблок дрозофилы может рассматриваться как способ их выявления и клонирования. Регуляторные механизмы, контролирующие развитие, могут оказаться более универсальными, чем ожидалось.

Молекулярная генетика развития делает свои первые шаги. Выявлены и клонированы гены, определяющие ключевые стадии развития. Начаты исследования особенностей пространственного распределения транскриптов генов и их продуктов в развитии эмбриона. Изучаются регуляторные элементы этих генов, в том числе те, от которых зависят закономерности пространственной экспрессии. Выявление гомеоблока, входящего в состав разных генов развития, позволило наметить подходы к исследованию генов, контролирующих развитие у разных организмов. Обнаружены белки, кодируемые этими генами и взаимодействующие с ДНК. Однако решение сложной задачи расшифровки законов, по которым строится трехмерная структура организма, еще впереди.

Развитие многоклеточных эукариотических организмов основано на способности клеток передавать в ряду поколений активное или, наоборот, репрессированное состояние гена. Наследование состояния гена приводит в конечном итоге к образованию дифференцированной ткани, состоящей из клеток, в которых лишь небольшая часть генов активирована на фоне репрессии основной части генома. Исследование молекулярных механизмов, обеспечивающих наследование активного или неактивного состояния гена в ряду клеточных поколений, представляется чрезвычайно важным. По-видимому, в основе этих механизмов лежат не только программированные взаимодействия белков и ДНК, обеспечивающие наследуемую локальную организацию хроматина, но и процессы метилирования ДНК. Метилирование можно рассматривать как особый механизм контроля транскрипции, существующий наряду с механизмами, основанными на взаимодействиях между цис-действующими регуляторными элементами и факторами транскрипции.

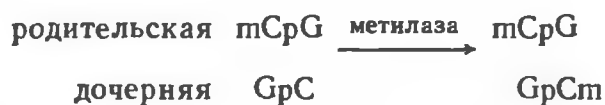
6. МЕТИЛИРОВАНИЕ

КАК СПОСОБ КОНТРОЛЯ АКТИВНОСТИ ГЕНОВ ЭУКАРИОТ

В ДНК эукариот около 5 % от общего содержания цитозина составляет 5-метилцитозин. Степень метилирования ДНК меняется в процессе развития, а также зависит от возраста организма и гормо-

нальных воздействий. Цитозин метилируется, если он находится в составе комплементарной последовательности $\begin{smallmatrix} mCpG \\ GpCm \end{smallmatrix}$. Особенно сильно метилирована ДНК позвоночных, где модифицированы от 60 до 90 % таких динуклеотидных последовательностей. Последовательности CpG сконцентрированы в отдельных районах генома, включающих несколько сотен нуклеотидных пар. Число таких районов у млекопитающих составляет около 30 000. Такие «островки» последовательностей CpG являются мишенями метилирования ДНК. «Рисунок» метилирования участков ДНК различается в разных тканях и на разных стадиях развития.

Метилирование может сопровождать процесс репликации ДНК, причем цитозин метилируется в дочерней нити ДНК только в том случае, если он метилирован и в родительской:



что и обеспечивает наследование либо метилированного, либо деметилированного состояния участка ДНК при последующих клеточных делениях.

Роль локального метилирования ДНК как фактора, контролирующего активность генов, подтверждается прежде всего реактивацией некоторых неактивных генов после частичного деметилирования. Метилирование можно ингибировать, добавляя к культурам клеток 5-азацитидин, подавляющий активность метилаз. В присутствии 5-азацитидина удается наблюдать резкое возрастание уровня экспрессии генов — в 10^5 — 10^6 раз, если, например, судить об этом по увеличению активности фермента. В таком случае эффект 5-азацитидина по силе своего действия сопоставим с эффектом мутационного события. Добавление 5-азацитидина иногда вызывает процессы дифференцировки, например образование мышечной ткани в культуре миобластов.

В генах, которые реактивируются, частичное деметилирование затрагивает как 5'-фланг, примыкающий к промотору, так и районы экзонов и интронов, или, наконец, 3'-фланги гена. Активное состояние гена характеризуется определенным «рисунком» распределения метилированных сайтов в районе гена, но не связано с полным деметилированием.

Репрессия генной активности наблюдается в результате прямого ферментативного метилирования генов. Метилированные гены, введенные в культуры клеток, сохраняют неактивное метилированное состояние в ряду поколений после многих актов репликации. Не ясно, является ли метилирование *in vivo* причиной инактивации генов или лишь «закрепляет» неактивное состояние, уже достигнутое, например, в результате предшествующего взаимодействия с белками.

После частичного деметилирования в результате воздействия 5-азацитидина активируются не все гены. Например, неактивные тканеспецифические гены глобинов не начинают транскрибироваться, несмотря на сильное деметилирование. Вероятно, деметилирование отдельных участков гена может быть необходимым, но не достаточным для обеспечения активной экспрессии гена.

Механизм действия метилирования не раскрыт. Модифицированная ДНК может оказывать влияние на локальную структуру в составе хромосомы. Вероятно, метилирование отдельных сайтов в составе гена меняет характер взаимодействия с белками и структуру хроматина. Действительно, сайты метилирования в отдельных исследованных генах совпадают с так называемыми гиперчувствительными к 'нуклеазам сайтами в составе хроматина, наличие которых отражает активное состояние гена или его готовность к активации (см. гл. XII). Метилирование может влиять и на структуру ДНК. Например, метилирование цитозина в составе синтетических полидезоксинуклеотидов с повторяющейся комплементарной последовательностью типа $d(CpG)_n \cdot d(GpC)_n$ способствует их переходу в Z-конформацию ДНК.

Роль метилирования несомненно очень велика при определении так называемого «родительского импринтинга» генов у млекопитающих. Известно, что роль отцовских и материнских генов в развитии не одинакова, степень и характер экспрессии генов может зависеть от того, получены они от отца или от матери. Было показано, что метилированные гены трансгенных мышей, полученные от матери, могут не экспрессироваться, тогда как эти же гены, полученные от отца и деметилированные в процессе гаметогенеза, способны к тканеспецифической экспрессии.

Процессы метилирования несомненно участвуют в инактивации одной из двух X-хромосом в клетках млекопитающих. Неактивное состояние одной из двух X-хромосом, возникающее в раннем развитии эмбриона, цитологически обнаруживается по наличию компактного гетерохроматического тельца Барра. Это неактивное состояние наследуется в клеточных поколениях, а реактивация X-хромосомы происходит при образовании герминальных клеток. Путем деметилирования с помощью 5-азацитидина также удавалось активировать гены неактивной X-хромосомы. По-видимому, инициация инактивации X-хромосомы обеспечивается взаимодействием со специфическими белками, а метилирование — это вторичный процесс, закрепляющий неактивное состояние X-хромосомы в последующих клеточных делениях.

Контроль генной активности с помощью метилирования, распространенный у позвоночных, вероятно, осуществляется не у всех эукариот. Например, *D. melanogaster* и нематода *Caenorhabditis elegans* — новый излюбленный объект молекулярной генетики, — вероятно, не используют метилирование ДНК для целей дифференциального контроля генной активности, поскольку в ДНК, выде-

ленной из этих организмов, количество 5-метилцитозина, не превышает по крайней мере одного остатка на несколько десятков тысяч нуклеотидных пар.

ЛИТЕРАТУРА

Ванюшин Б. Ф. Метилирование ДНК у эукариот — новый механизм регуляции экспрессии генов и клеточной дифференцировки // Усп. биол. химии. 1983. Т. 24. С. 170—193.

Геринг В. Молекулярные основы развития // В мире науки. 1985. № 12. С. 111—121.

Рэфф Р., Кофмен Т. Эмбрионы, гены и эволюция. — М.: Мир. 1986.

Adams P. L., Burdon R. H. Molecular biology of DNA methylation. Berlin Heidelberg — New York Springer. 1985.

Akam M. The molecular basis for metameric pattern in the *Drosophila* embryo // Development. 1987. V. 101. P. 1—22.

Cedar H. DNA methylation and gene activity // Cell, 1988, V. 53. P. 3—4.

DNA methylation: biochemistry and biological significance. Eds A. Razin, H. Cedar and Riggs NY — Springer Verlag. 1984.

Guarente L. Regulatory proteins in yeast // Ann. Rev. Genet. 1987. 21. P. 425—452.

Maniatis T., Goodbourn S., Fisher J. A. Regulation of inducible and tissue specific transcription. Science. 1987. V. 236. P. 1237—1244.

Palmiter R. D., Brinster R. L. Transgenic mice // Cell. 1985. V. 41. P. 343—345.

Глава XI

ПОДВИЖНЫЕ ГЕНЕТИЧЕСКИЕ ЭЛЕМЕНТЫ ГЕНОМА ЭУКАРИОТ

Существенную часть генома эукариот (10—20 %) составляют повторяющиеся последовательности ДНК (см. гл. X), которые в основном представлены разными типами подвижных (мобильных) элементов. Критерием для отнесения фрагментов генома к числу подвижных часто служит лишь локализация по их флангам коротких прямых повторов, обычно включающих несколько нуклеотидных пар ДНК, что отражает молекулярные механизмы акта их внедрения в ДНК-мишень (см. гл. IV). Функциональная роль подвижных элементов не ясна, но, во всяком случае, они в значительной степени определяют, как и у прокариот, изменчивость генома и, следовательно, могут играть большую роль в эволюции генома. Роль мобильных элементов в мутационном процессе, включая образование делеций и инверсий, уже была рассмотрена на примере прокариот (см. гл. VI).

Многие спонтанные мутации эукариот, например у дрозофилы, также обусловлены внедрением мобильных элементов. Случаи внедрения и перемещения (транспозиции) таких мобильных элементов представляют собой достаточно редкие события (10^{-4} — 10^{-5} в расчете на ген в одном поколении). Положение мобиль-

ного элемента может сохраняться неизменным в данном сайте генома, в таком случае их можно рассматривать как мобильные только в эволюционном масштабе времени. Однако в клетке могут возникать условия, определяемые, с одной стороны, внутренними генетическими факторами, а с другой — внешней средой (например, изменением температуры), когда может резко увеличиваться частота транспозиций некоторых типов мобильных элементов. Наблюдаемые в отдельных случаях множественные транспозиции мобильных элементов будут приводить к существенным геномным перестройкам, которые могут сказываться на эволюционной судьбе организмов.

Выделяют ряд классов подвижных элементов эукариот на основании различий их молекулярной структуры и способности к перемещениям.

1. ЭЛЕМЕНТЫ ГЕНОМА, ПРЕДСТАВЛЯЮЩИЕ СОБОЙ ПРОДУКТЫ ОБРАТНОЙ ТРАНСКРИПЦИИ КЛЕТОЧНЫХ РНК (РЕТРОПОЗОНЫ)

Детальное исследование молекулярной организации генома высших эукариот, особенно млекопитающих, показало, что существенная часть генома, около 10 % общей массы ДНК, образовалась в результате интеграции в геном фрагментов ДНК, синтезированных на РНК-матрицах в результате обратной транскрипции (рис. 118, а). Впервые подобный процесс был описан при исследовании ретровирусов, в геноме которых имеется ген, кодирующий обратную транскриптазу (ревертазу) (см. гл. XIII). В геноме млекопитающих, птиц, амфибий и насекомых обнаруживаются *ретропозоны*, представляющие собой внедрившиеся в геном ДНК-копии, синтезированные на разных типах клеточных РНК как на матрицах. Молекулярные механизмы «ретропозиции» не изучены, остается не установленным источник клеточной обратной транскриптазы. Не ясно, что служит затравкой для ревертазы; возможно, это шпилька на 3'-конце РНК, образующаяся в результате комплементарных взаимодействий. Как будет видно, структура ретропозонов позволяет с уверенностью говорить об участии обратной транскрипции в процессе их образования. Таким образом, наряду с переносом информации от ДНК к РНК осуществляется и обратный процесс — возвращение ее в геном в виде ретропозонов. У млекопитающих ретропозоны составляют более 10 % ДНК; следовательно, мощность встречного потока информации от РНК к ДНК может быть существенной, по крайней мере при оценке его во временном эволюционном масштабе. Различают разные типы ретропозонов.

В геномах млекопитающих, иногда вблизи генов с интронами, активно транскрибируемых РНК-полимеразой II, обнаружива-

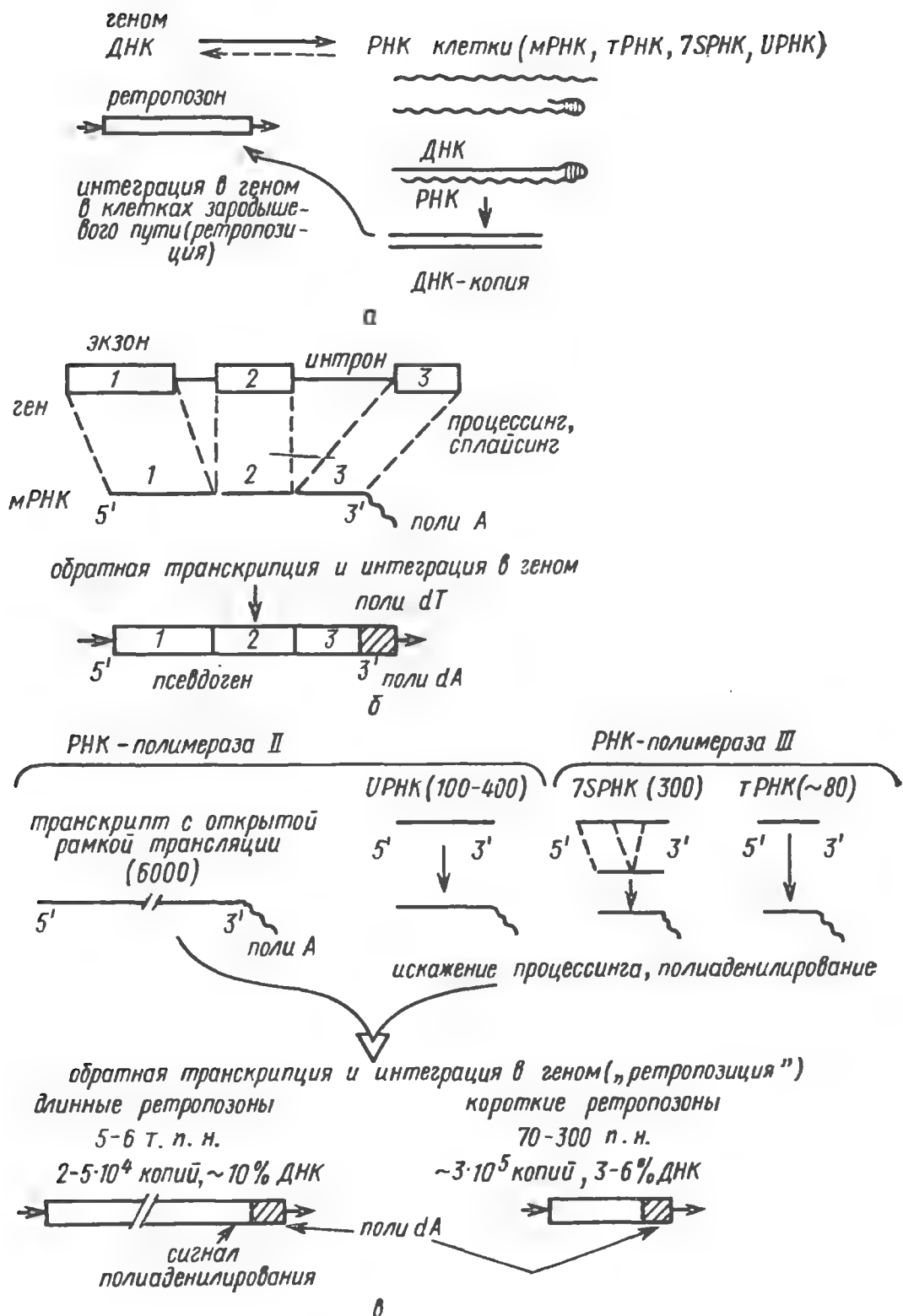


Рис. 118. Формирование элементов генома при участии процессов обратной транскрипции

Стрелки — прямые повторы по краям перемещающегося элемента («ретропозона»); а — общая схема (пунктирная стрелка указывает на существование «потока информации», направленного от РНК к геномной ДНК); б — образование псевдогенов на «процессированной» мРНК как на матрице; в — многокопийные короткие и длинные повторы (ретропозоны) в геноме млекопитающих, в скобках указаны размеры РНК в нуклеотидах

ются так называемые *псевдогены*, представляющие собой испорченные копии этих генов, лишенные интронов и обладающие последовательностью поли dA-поли dT на 3'-фланге. Следовательно, матрицей при их копировании послужила «процессированная» полиаденилированная мРНК (рис. 118, б). Псевдогены получили такое название, поскольку рамки трансляции таких «генов» часто испорчены стоп-кодонами и делециями. У человека, например, обнаружены псевдогены, соответствующие генам глобинов, триозофосфатизомеразы, рибосомального белка L32; у мыши — генам цитохрома с, рибосомальных белков L7 и L18, глицеральдегидтрифосфатдегидрогеназы; у дрозофилы — гену алкогольдегидрогеназы.

Псевдогены представлены в геноме разным числом копий — от одной до нескольких десятков. Они, как и проретровирусы (см. гл. XIII), ограничены короткими прямыми повторами, представляющими собой дубликацию геномной последовательности в области сайта внедрения ДНК-копии, образованной на мРНК. Иногда псевдоген не содержит 5'-конца гена, в таком случае можно предполагать, что обратная транскрипция не была доведена до 5'-конца молекулы мРНК. Поскольку необходимые промоторные элементы РНК-полимеразы II локализованы перед стартом транскрипции, внедрение ДНК-копии правильно инициированной РНК приведет к образованию ретропозона, лишенного собственного промотора, и, следовательно, возможности транскрипции. Однако при внедрении ретропозона в район чужого промотора транскрипция может осуществляться, но она будет управляться уже другими регуляторными элементами. Если ретропозон правильно процессирован и ДНК-копия содержит открытую рамку трансляции, то возможна экспрессия ретропозона. Такие ретропозоны были обнаружены, их иногда называют ретрогенами, подразумевая возможность их функционирования в клетке. В состав ретрогена попадет и собственный промотор гена, если матрицей будет служить транскрипт, инициированный за несколько сотен нуклеотидных пар перед *кэп*-сайтом. Вероятно, это произошло при образовании «полупроцессированного» гена препроинсулина (предшественника инсулина) у грызунов. Мыши и крысы обладают двумя генами препроинсулина. Ген II содержит большой и малый интрон, а ген I, ограниченный прямыми повторами, — только малый; большой интрон делетирован с сохранением открытой рамки трансляции. По-видимому, транскрипт, послуживший матрицей для обратной транскриптазы, был инициирован перед нормальным сайтом инициации и включил район промотора. Наличие гена I препроинсулина, возникшего на определенной стадии эволюции и активно транскрибируемого, возможно, обеспечило животным особые селективные преимущества. Таким образом, внехромосомные цитоплазматические молекулы РНК становятся снова родоначальницами геномной последовательности, причем характер их процессинга может определять особенности структуры новообра-

зованного гена. Случаи образования процессированных генов — это достаточно редкие события, приводящие к изменению структуры генома лишь в эволюционном масштабе времени.

Псевдогены, образованные на мРНК, предназначенных для кодирования известных белков, составляют небольшую часть геномных ретропозонов. Основная масса повторяющихся элементов позвоночных, представленных огромным числом копий — десятками или сотнями тысяч, также образовалась в результате ретропозиции ДНК-копий клеточных РНК, однако матрицами для них служили полиаденилированные РНК, кодирующие белки неизвестной природы, а также аномально процессированные клеточные транскрипты тРНК, 7SPHK и UPHK (см. рис. 118, в). Особенно богаты такими повторами геномы высших эукариот — млекопитающих. У них обнаружены рассеянные по геному короткие (70—300 п. н.) и длинные (6—7 тыс. п. н.) повторы. Те и другие несут на условном 3'-конце последовательность полидА и ограничены прямыми повторами (7—20 п. н.) (рис. 118, в), представляющими собой дубликацию сайта внедрения ретропозона. Внутри семейств как длинных, так и коротких повторов обнаруживается структурная неоднородность. Отдельные копии часто укорочены с 5'-конца, имеют делеции, вставки и нуклеотидные замены. Семейства таких повторов часто получали свои названия по рестриктазам, расщепляющим эти последовательности (например, короткие Alu-повторы или длинные Kpn-повторы в геноме человека).

Число копий длинных повторов составляет у млекопитающих $\sim 3 \cdot 10^4$, т. е. около 10 % массы ДНК. В длинных повторах обнаруживается одна или несколько достаточно протяженных открытых рамок трансляции, где выявляется заметная гомология с нуклеотидными последовательностями, кодирующими ревертазу у ретровирусов. Гомология нуклеотидных последовательностей открытых рамок трансляции прослеживается у разных видов, однако за пределами открытой рамки трансляции гомология теряется. Полагают, что родоначальницей длинных повторов млекопитающих послужила фракция полиаденилированных РНК, кодирующих белки, обладающие ревертазной активностью. Соответствующие полиаденилированные транскрипты размером 6000 нуклеотидов обнаруживаются в недифференцированных эмбриональных клетках млекопитающих. Не исключено, что в результате трансляции таких транскриптов образуется клеточная ревертаза.

Каждый вид млекопитающих обладает собственным главным семейством длинных повторов, отличающим его, особенно вне области открытой рамки трансляции, от других видов. Поэтому предполагается, что разные семейства длинных повторов возникли сравнительно недавно, уже после эволюционной дивергенции видов млекопитающих. У особей — родоначальников каждого вида, вероятно, существовала своя предковая последовательность, размножение которой привело к образованию семейства длинных по-

второв, часто обнаруживаемых в промежутках между генами.

Короткие повторы, по-видимому, возникли в результате ретропозиции ДНК-копий, образованных на частично процессированных тРНК и 7S РНК. Последняя выполняет важную клеточную функцию, она входит в состав рибонуклеопротеидных частиц, узнающих сигнальный пептид на N-конце новообразованного секретируемого белка. Частицы участвуют в биосинтезе секретируемых белков, обеспечивая прикрепление рибосомы к мембранам эндоплазматической сети. Анализ нуклеотидной последовательности коротких повторов показывает, что матрицей для их образования послужили 7S РНК, тРНК и, наконец, УРНК, подвергшиеся процессингу и «ошибочному» полиаденилированию. Возможно, образование шпильки на 3'-конце молекул обеспечило затравку (праймер) для обратной транскриптазы.

Геном млекопитающих содержит несколько разных семейств коротких повторов. Короткие повторы у птиц и амфибий изучены значительно хуже. Число копий коротких повторов, например наиболее изученных повторов Alu-семейства у человека, составляет $3 \cdot 10^5$, что соответствует 5—6 % массы ДНК клетки. Такие повторы рассеяны по геному и получили название вездесущих. Повторы Alu могут находиться в интронах, на 5'-флангах генов и, наконец, в составе 3'-нетранслируемого участка мРНК. Нуклеотидная последовательность Alu-повтора гомологична последовательности отдельных участков 7S РНК. Структура 7S РНК достаточно консервативна у позвоночных, а гомологии в нуклеотидной последовательности прослеживаются и с 7S РНК насекомых. Поэтому семейства коротких повторов, присутствующие у разных видов, предшественником которых служила 7S РНК, также могут обладать достаточной гомологией. В то же время семейства коротких повторов, как и длинных, характеризуются видоспецифичностью, обусловленной амплификацией той или иной копии клеточных РНК, которые к тому же могли быть по-разному модифицированы в результате процессинга. Локализация ретропозонов, внедрившихся в отдельные сайты генома у предков млекопитающих, может, по крайней мере, частично сохраняться в процессе дальнейшей эволюции. Например, места локализации Alu-подобного семейства в межгенных промежутках кластера глобиновых генов оказались достаточно сходными у мышей и приматов.

Функция коротких и длинных повторов не ясна. Они транскрибируются и обнаруживаются в составе протяженных ядерных транскриптов, которые еще не подверглись процессингу. Короткие повторы содержат потенциальные промоторы для полимеразы III, поскольку они образовались на РНК-матрицах, синтезированных с помощью РНК-полимеразы III, использующей для инициации транскрипции участки внутри транскрибируемого гена (см. гл. X).

Однако непонятно, используются ли в клетке эти промоторы коротких повторов. Между особями одного вида иногда наблюдаются различия (полиморфизм) в структуре генов и их ближайшего окружения, обусловленные внедрением коротких или длинных повторов. Однако вопрос о возможной селективной ценности (например, о преимуществах при экспрессии гена) этих различий остается открытым.

2. ПОДВИЖНЫЕ ЭЛЕМЕНТЫ С ДЛИННЫМИ КОНЦЕВЫМИ ПОВТОРАМИ (РЕТРОТРАНСПОЗОНЫ)

Другой большой класс подвижных элементов составляют *ретротранспозоны*, сходные по своей структуре с проретровирусами (см. гл. XIII), которые внедряются в геном, используя механизмы обратной транскрипции. Эти элементы содержат «тело» размером 5—8 т. п. н., ограниченное прямыми длинными концевыми повторами (ДКП, или LTR, англ. long terminal repeats), обычно содержащими по 300—400 п. н. (рис. 119, а). Число копий таких элементов, принадлежащих к одному семейству, достаточно постоянно для вида, но варьирует от нескольких копий до сотен тысяч копий в зависимости от типа ретротранспозона. Ретротранспозоны, относящиеся к разным семействам, не имеют протяженных гомологий нуклеотидных последовательностей ни в ДКП, ни в теле. В составе «тела» элемента обнаруживаются открытые рамки трансляции. Внутри них показана локальная гомология тем районам геномов проретровирусов, которые кодируют обратную транскриптазу и нуклеазу (интегразу). Иногда обнаруживается и открытая рамка для белка, гомологичного по аминокислотной последовательности белку оболочки ретровирусов. Разные копии одного семейства могут различаться наличием вставок и делеций как в ДКП, так и в самом теле элемента. По флангам ДКП находятся отражающие дупликацию ДНК мишени короткие прямые повторы, длина которых (несколько нуклеотидных пар) характерна для каждого элемента. Элементы подобной структуры обнаружены у дрожжей, дрозофилы, позвоночных и растений. Ретротранспозоны обычно ведут себя как стабильно наследуемые гены, однако определенные воздействия окружающей среды могут индуцировать их перемещение, как это наблюдается, например, в клетках дрожжей.

Один из способов перемещения требует прежде всего образования РНК-матрицы, которая копируется при участии обратной транскриптазы (рис. 118). Это было экспериментально доказано для ретротранспозонов дрожжей и дрозофилы. Ретротранспозоны транскрибируются с помощью РНК-полимеразы II. В составе ДКП имеются сайты инициации транскрипции и сигналы полиаденилирования. ДКП могут служить активными промоторами. Транскрипция начинается в одном ДКП (условно левом, 5'-ДКП)

и заканчивается в другом (правом, 3'-ДКП). По-видимому, асимметрия самого тела элемента и различия в геномном окружении двух ДКП определяют выбор инициации транскрипции элемента только с одного из двух одинаковых ДКП. По-видимому, элементы, которые кодируют белки, участвующие в «ретропозиции», т. е. ревертазу и др., обеспечивают наличие транс-действующих факторов, необходимых для перемещения сходных, но дефектных элементов данного семейства, не способных кодировать эти белки.

Транскрипция ретротранспозонов далеко не всегда сопровождается образованием ДНК-копии и ретропозицией. В ряде случаев, например на ранних эмбриональных стадиях развития нормальных мышей, продукты транскрипции отдельных семейств ретротранспозонов находятся в составе вирусоподобных цитопламати-

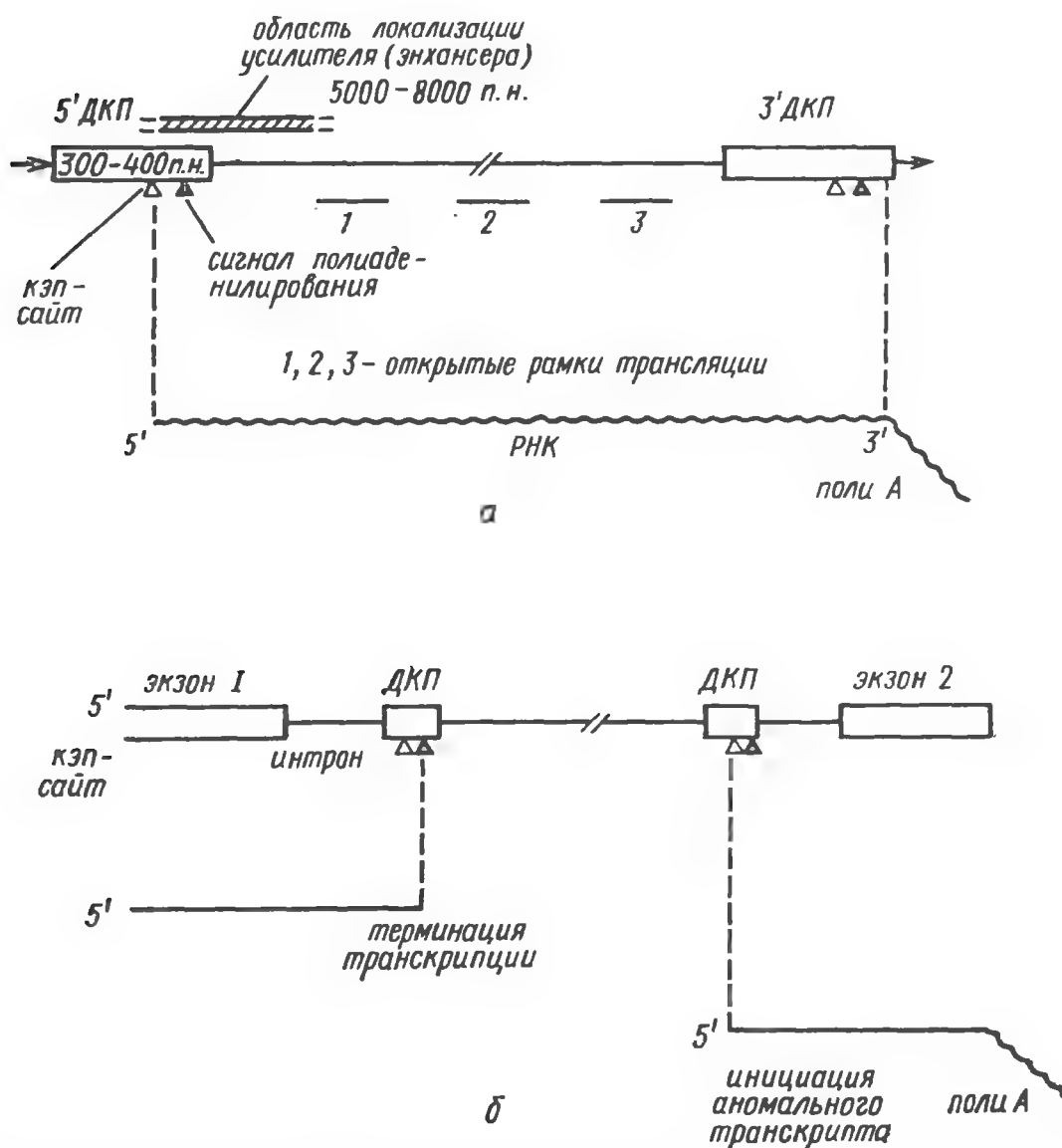
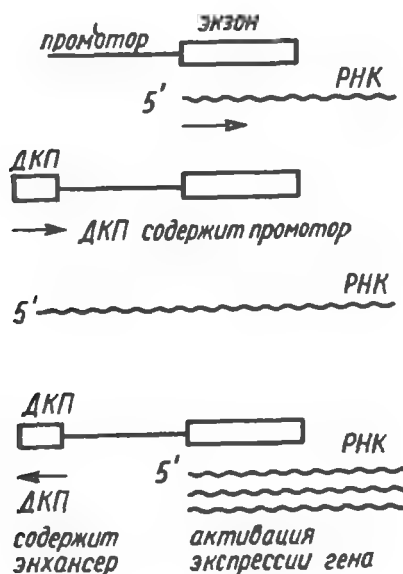


Рис. 119. Сходные по структуре с проретровирусами ретротранспозоны эука-
а — общая структура; б — нарушения транскрипции, вызванные внедрением ретротранспо-
направление транскрипции); г — вырезание и встраивание в результате гомологичной реком-

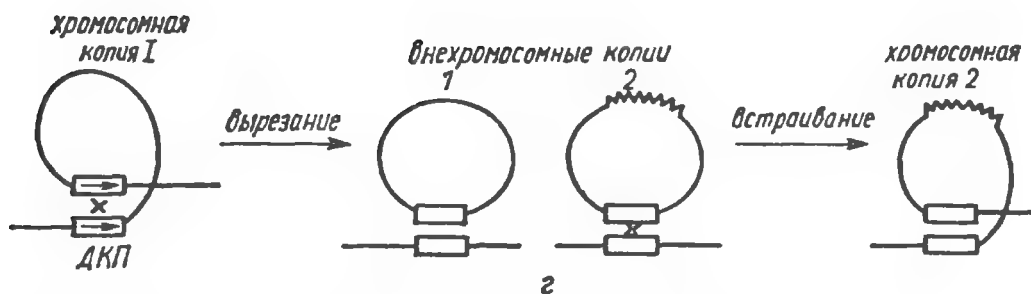
ческих частиц. Эти частицы не выходят из клетки и не способны заражать другие клетки. Остается загадкой, какую роль они могут играть в жизнедеятельности клетки и в развитии организма.

Элементы, ограниченные ДКП, способны также перемещаться, используя рекомбинационные механизмы. В результате гомологичной рекомбинации между ДКП элемент вырезается из хромосомы, но в составе генома сохраняется один ДКП, присутствие которого сможет обеспечить повторную интеграцию, например, какой-либо другой копии из числа представителей данного гетерогенного семейства (рис. 119, з).

С одной стороны, имеются свидетельства о достаточно древних случаях внедрений ретротранспозонов, затем сохранивших неизменным свою локализацию. Так, в промежутке между генами эмбриональных глобинов γ^A и γ^G обнаружен один и тот же элемент у людей, относящихся к разным расам. С другой стороны,



Б



риот, ограниченные длинными концевыми повторами (ДКП):

зона в нитрон; σ — усиление экспрессии гена при внедрении в 5'-район стрелки указывают бицинации в районах ДКП

имеются данные о внезапных перемещениях сразу многих копий элементов, наблюдаемых в процессе эксперимента. Внезапные перемещения элементов, ограниченных ДКП, описаны у дрожжей и дрозофилы, но причины их плохо поняты. Возможно, первопричиной размножения копий является их усиленная транскрипция. Оказалось, что резкая активация транскрипций лишь одной полной копии ретротранспозона дрожжей сопровождается размножением не только этой копии, но и многих других, в том числе, дефектных, копий. Это удалось показать, создав искусственную генетическую конструкцию, в составе которой элемент был слит с промотором гена *GAL1*, который индуцируется галактозой. Добавление к дрожжевым клеткам галактозы резко активировало транскрипцию этой копии мобильного элемента, в результате чего в клетках образовывались белковые продукты, необходимые для ретропозиции ряда других копий. Этот пример показывает, как условия внешней среды могут оказывать влияние на процессы ретропозиции.

Внедрение мобильного элемента внутрь гена или около гена вызывает разные эффекты. Во многих случаях происходит инактивация гена, например нарушается образование нормальных транскриптов в результате терминации вблизи сайтов полиаденилирования в одних ДКП или, наоборот, инициации в других ДКП (рис. 119, б). При интеграции в район промотора на 5'-фланге гена мобильный элемент может резко активировать экспрессию гена, обеспечивая транскрипцию с собственного промотора. Однако активирующее влияние элемента может наблюдаться, если направления транскрипции в ДКП и в гене противоположны. Возможно, активация транскрипции и экспрессии гена осуществляется в таком случае благодаря воздействию энхансеров, привносимых элементом (рис. 119, в). Действительно, в составе ДКП или «тела» ряда мобильных элементов находятся нуклеотидные последовательности, ведущие себя как энхансеры, т. е. действующие независимо от ориентации по направлению к транскрипции гена (см. гл. X).

Описанные случаи внедрения элемента сопровождаются мутациями с самыми разными фенотипическими проявлениями, обусловленными подавлением образования или, наоборот, гиперпродукцией белка. Можно наблюдать полную или частичную реверсию мутаций к норме, вызванную вырезанием мобильного элемента при сохранении в составе хромосомы только одного ДКП. Перемещение мобильных элементов по геному могут способствовать распространению регуляторных сигналов (сайтов инициации транскрипции, сигналов полиаденилирования или энхансеров). Роль мобильных элементов в эволюции систем регуляции может быть значительной, если принять во внимание, что геном эукариот кодирует трансдействующие белковые факторы, способные специфически регулировать инициацию транскрипции в районе ДКП.

3. МОБИЛЬНЫЕ ЭЛЕМЕНТЫ, ОГРАНИЧЕННЫЕ ИНВЕРТИРОВАННЫМИ ПОВТОРАМИ

Эти элементы ограничены инвертированными повторами, как и некоторые транспозоны прокариот (рис. 120, а). Примерами их могут служить Р-элемент дрозофилы и Ас-элемент кукурузы. Их рассмотрение показывает, какую важную биологическую роль они могут играть в определении геномной изменчивости как в соматических клетках, так и в клетках зародышевого пути. В геномах дрозофилы и кукурузы насчитывается по 30—50 копий таких элементов. Полные копии содержат открытые рамки трансляции, кодирующие транспозазу. Часть копий дефектна, содержит внутренние делеции, и их транспозиции, по-видимому, осуществляются с помощью транспозазы благодаря функционированию полных копий. В «теле» Р-элемента обнаружены четыре открытые рамки трансляции, которые в результате сплайсинга объединяются в составе мРНК, кодирующей транспозазу. Транскрипт Ас-элемента также захватывает все тело, после удаления четырех интронов образуется мРНК для транспозазы.

Р-элемент содержится только в некоторых линиях *D. melanogaster*.

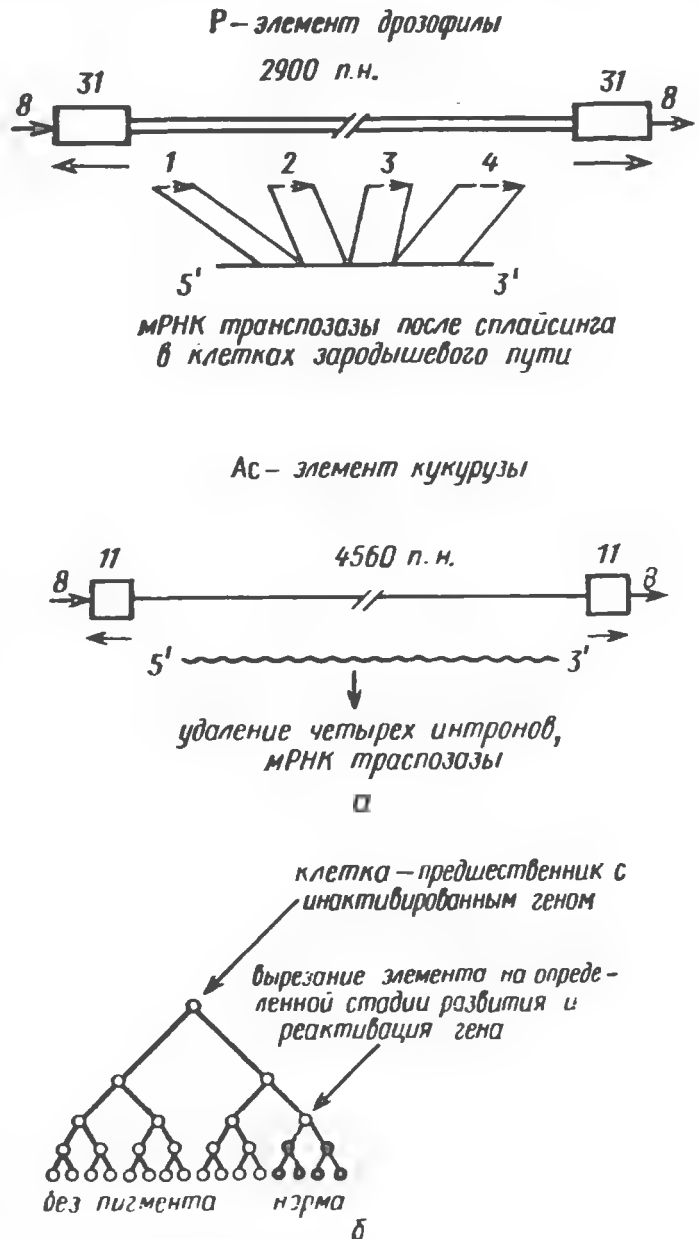


Рис. 120. Мобильные элементы эукариот с концевыми инвертированными повторами:

а — сплошные стрелки — положение прямых (результат дупликации в районе сайта внедрения) и инвертированных повторов (цифрами указаны размеры в нуклеотидах; пунктирные стрелки обозначают открытые рамки трансляции); б — клональное наследование инактивированного состояния гена и его реактивация при вырезании элемента в соматических клетках растений на определенных стадиях развития

Скрещивание самок без Р-элемента с самцами, несущими Р-элементы, приводит у гибридов к транспозициям Р-элемента, которые наблюдаются только в клетках зародышевого пути. В потомстве таких гибридов обнаруживается достаточно много мутаций, вызванных внедрением элемента. Эти мутации часто приводят к стерильности потомства. Поэтому линии с Р-элементом и без него выглядят как репродуктивно изолированные, по крайней мере частично. Биологическая изоляция играет огромную роль в процессе эволюции. В этом случае она объясняется на молекулярном уровне: изоляция линий вызвана активацией транспозиций Р-элемента, присутствующего в одной из них. Механизм активации транспозиций не расшифрован, однако выяснена причина, почему транспозиции Р-элемента ограничены зародышевыми клетками. Оказалось, что только в клетках—предшественниках гамет — осуществляется такой ход сплайсинга транскрипта Р-элемента, который приведет к образованию непрерывной открытой рамки трансляции, кодирующей транспозазу (рис. 120, а). Ограничение транспозиции зародышевыми клетками, по-видимому, имеет определенный смысл, поскольку обеспечивает выживание особей, несущих гаметы, в которых произошли геномные перестройки вследствие транспозиции Р-элемента. Подобный «геномный шок», сопровождающийся высокой частотой мутагенеза, может обеспечить большую степень геномной изменчивости, которая послужит материалом для отбора в процессе эволюции.

Если транспозиции Р-элемента дрозофилы ограничены зародышевыми клетками, то перемещения Ас-элемента происходят и в соматических клетках у кукурузы. За перемещением таких элементов можно следить по распределению фенотипически нормальных и мутантных участков ткани — например, лишенных пигмента вследствие инактивации гена, определяющего пигментацию. Потомство клетки, содержащей только инактивированный ген(ы), также будет лишено пигмента. Вырезание мобильного элемента приводит к реактивации гена. Чем раньше оно произойдет в развитии мутантной непигментированной ткани, тем обширнее будет окрашенный участок, поскольку клетки наследуют активное состояние гена (рис. 120, б). Наблюдая подобные явления, Мак-Клинток сделала вывод о регуляторной функции перемещающихся элементов, назвав их контролирующими. Оказалось, что вырезание этих элементов происходит только в определенных тканях и в течение ограниченного периода развития растения.

Вырезание Ас-подобных элементов у растений может быть неточным — элемент оставляет «отпечаток» своего присутствия внутри гена, образуя, например, дупликацию или вставку нескольких нуклеотидов в районе внедрения. В результате могут изменяться свойства белка или характер экспрессии гена, если элемент соответственно побывал в экзоне или в регуляторном районе гена.

В целом подвижные элементы эукариот представляют собой

чрезвычайно разнородную популяцию и им можно приписывать самые разные функции в процессе функционирования генома. Однако существует представление о том, что описанные подвижные элементы не несут никакой функции, т. е., как правило, не влияют на фенотип организма и размножаются в геноме лишь благодаря особенностям своей структуры, в результате чего они постепенно заселяют геном. Предполагается, что они составляют фракцию так называемой эгоистичной (англ. selfish) ДНК, размножение которой ограничивается естественным отбором. Отбор устраняет те случаи, когда внедрение эгоистичной ДНК сопровождается вредными последствиями. Недостаток знаний о мобильных элементах эукариотического генома не позволяет ни отвергнуть, ни принять целиком эту концепцию. Представляется вероятным, что на определенных стадиях эволюции отдельные элементы поначалу эгоистичной ДНК могут использоваться для целей, которые в дальнейшем обеспечат селективные преимущества организму.

ЛИТЕРАТУРА

- Вармус Х. Обратная транскрипция// В мире науки. 1987. № 11. С. 22—29.
Ильин Ю. В. Мобильные диспергированные гены эукариот// Итоги науки и техники. Сер. мол. биол. М.: ВИНТИ. 1982. Т. 18. С. 5—48.
Rogers J. H. The origin and evolution of retroposons// Int. Rev. of Cytol. 1985. V. 93. P. 187—279.
Doring H.—P., Starlinger P. Molecylar genetics of transposable elements in plants// Ann. Rev. Genet. 1986. V. 20. P. 175—200.
Vanin E. Processed pseudogenes. Characteristics and evolution// Ann. Rev. Genet. 1985. V. 19. P. 253—272.

Глава XII

СТРУКТУРА ХРОМАТИНА

Наследственный аппарат эукариотических клеток существенно отличается от прокариотических хромосом. Наиболее очевидное отличие — огромное количество ДНК в эукариотических клетках. Например, гаплоидный геном человека состоит из $3 \cdot 10^9$ пар оснований (п. о.), тогда как геном *E. coli* включает всего 10^6 п. о. Кроме того, геном эукариот разделен на несколько хромосом, которые претерпевают характерные циклы конденсаций и деконденсаций в ходе деления клеток. Наконец, в клетках эукариот больше генов и их регуляция значительно сложнее, чем у прокариот.

Структурная организация генетического материала в эукариотических клетках обусловлена следующими его биологическими функциями. Во-первых, ДНК в ядрах конденсируется примерно в 10^4 раз к моменту митоза, а затем быстро деконденсируется,

чтобы стать доступной для транскрипции, причем процесс конденсации должен быть упорядоченным, чтобы ДНК не запутывалась. Конденсация ДНК происходит в несколько этапов. Первый уровень конденсации — образование повторяющихся единиц, *нуклеосом*, состоящих из белковой глобулы и обвивающей ее спиральной ДНК. Из нуклеосом образуется первичная *хроматиновая фибрилла* толщиной около 10 нм. Она участвует в новом процессе спирализации с образованием фибриллы толщиной 30 нм, которая претерпевает дальнейшую компактизацию вплоть до образования хромосом. Важнейшую роль в образовании 30-нм фибриллы играет гистон H1.

Во-вторых, со сложными структурными перестройками хроматина связана репликация ДНК. Как будет обсуждаться ниже, в момент прохождения репликационной вилки ДНК сбрасывает гистоны и почти сразу после этого нуклеосомы реконструируются на ДНК. Следовательно, должны существовать механизмы, которые обеспечивают транспорт новосинтезированных гистонов из цитоплазмы в ядро и сборку нуклеосом на ДНК.

В-третьих, ДНК транскрибируется, и транскрипция различных генов тонко регулируется, в частности, на различных стадиях клеточного цикла и в процессе дифференцировки многоклеточных организмов. Гистоны, связанные с ДНК, влияют на этот процесс, они должны или удаляться с ДНК в момент транскрипции, или каким-то иным способом пропускать РНК-полимеразу. Механизмы узнавания белками определенных последовательностей ДНК у эукариот изучены в гораздо меньшей степени, чем у прокариот. Возможно, у эукариот важную роль в этом процессе играют белок-белковые взаимодействия. Многие эукариотические гены подчиняются нескольким различным регуляторным сигналам, поэтому их система регуляции весьма сложна и наверняка включает несколько белков.

Структуру эукариотических хромосом (*хроматина*) изучают с помощью различных подходов, в первую очередь биохимических и электронно-микроскопических. Биохимические исследования обычно основаны на выделении препарата ядер. Ядро — самая крупная и тяжелая (по плотности) органелла клеток. Препарат ядер довольно легко получить. Для этого ткань или клетки разрушают и центрифугируют, а затем очищают ядра, пропуская их через плотный раствор сахарозы с помощью повторного центрифугирования. Полученные ядра стабилизируют в процессе выделения двухвалентными катионами (Ca^{2+} или Mg^{2+}), полиаминами, а также 0,15 М NaCl, т. е. близкой к физиологической ионной силой. Такой препарат ядер сохраняет многие прижизненные свойства, в том числе способность синтезировать РНК и ДНК.

Если понизить ионную силу раствора и связать или отмыть двухвалентные катионы, то ядра «вскрываются», т. е. их содержимое превращается в 10-нм фибриллы. При этом теряются многие

компоненты ядер, а ДНК остается в виде прочного комплекса с гистонами. Соотношение гистоны/ДНК по массе в ядрах и в 10-нм фибриллах всегда близко к единице.

1. ГИСТОНЫ

Имеется пять основных типов молекул гистонов. Четыре из них — нуклеосомные гистоны H2A, H2B, H3 и H4 — сравнительно небольшие белки с $M_r = 10\,000 \div 15\,000$. Они очень богаты положительно заряженными аминокислотами — лизином и аргинином и сходны по структуре. Основная масса положительно заряженных аминокислот сосредоточена в N-концевых областях и в меньшей степени в С-концевых областях, тогда как центральная часть богата гидрофобными остатками. Заряженные концы молекулы находятся в не упорядоченной конформации, тогда как середина организована в основном в α -спирали и имеет глобулярную конформацию (рис. 121).

Гистон H1 существенно отличается от других гистонов. Он не входит в состав минимальных нуклеосом (см. раздел 4 этой главы) и участвует в организации 30-нм фибриллы хроматина. Его молекулярная масса превышает 20 000. Положительно заряженные аминокислотные остатки H1, главным образом лизины, находятся в основном в С-конце молекулы и в меньшей степени в N-концевой части. Центральная область N-концевой половины молекулы богата гидрофобными остатками и образует глобулу. H1 обладает выраженной доменной структурой, мягкое расщепление трипсином легко делит его на глобулу и «хвост». Помимо лизинов «хвост» богат остатками пролина и глицина и имеет неупорядоченную конформацию.

Взаимодействие ДНК с гистонами происходит главным образом путем образования ионных связей отрицательно заряженных фосфатов с положительно заряженными аминокислотными остатками гистонов. Поэтому комплексы ДНК с гистонами легко диссоциируют при высокой концентрации соли, например в 1,5—2 М NaCl. В то же время центральные гидрофобные участки гистонов взаимодействуют между собой.

Гистоны присутствуют во всех ядрах эукариотических клеток и отличаются чрезвычайно высокой эволюционной консервативностью. Так, например, при сравнении наиболее консервативного гистона H4 быка и гороха наблюдаются лишь две консервативные замены: там, где в белке быка находятся валин и лизин, в гистоне гороха расположены изолейцин и аргинин. Гены, кодирующие гистоны, транскрибируются РНК-полимеразой II. В эукариотическом геноме обычно содержатся от нескольких десятков до нескольких сотен tandemно расположенных генов гистонов (рис. 122). Они располагаются одинаковыми группами по пять разных генов, причем каждый ген транскрибируется по отдельности. Иногда

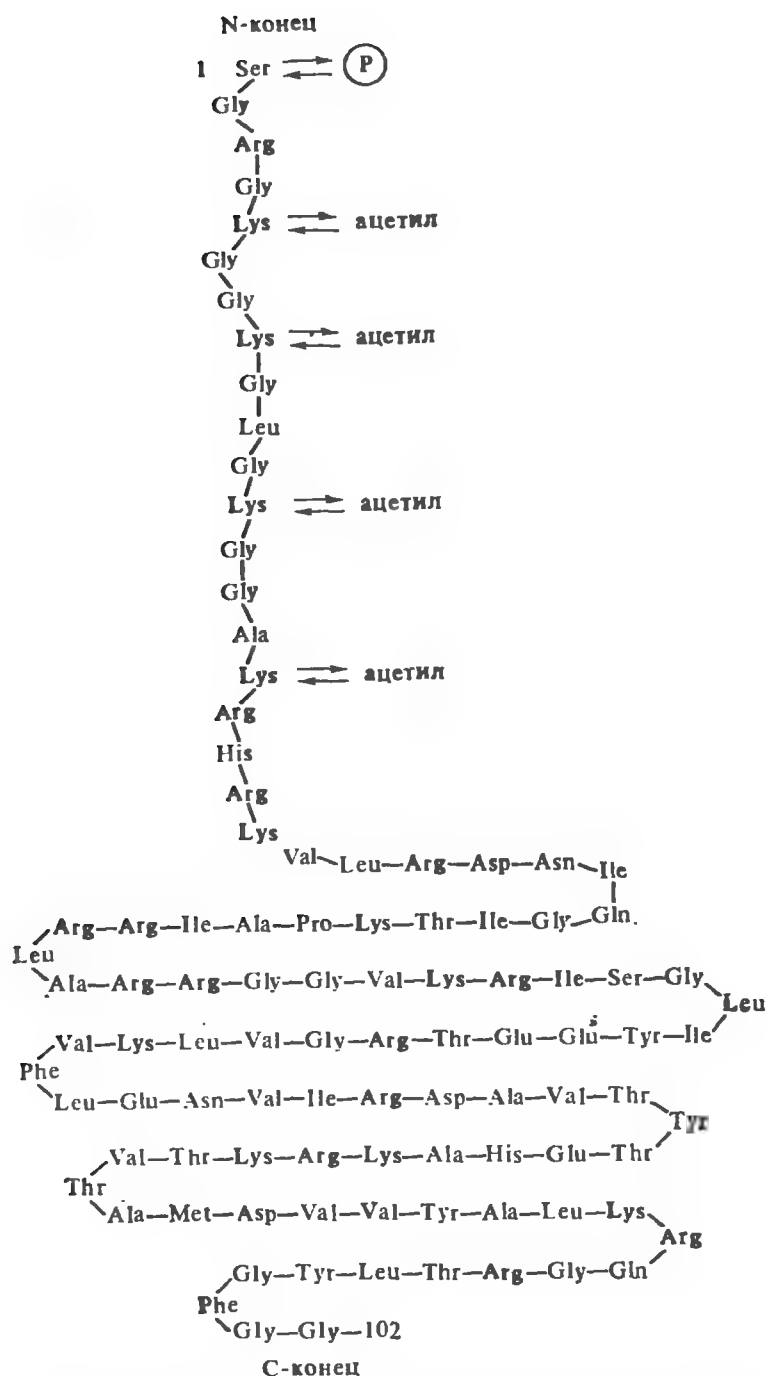


Рис. 121. Аминокислотная последовательность гистона H4

Показаны разупорядоченная N-концевая часть и глобулярный домен; указаны места ацетилирования и участок фосфорилирования (P); жирный шрифт — положительно заряженные аминокислоты

геном одного организма содержит несколько вариантов генов одного гистона, незначительно различающихся по аминокислотной последовательности. Наибольшие вариации наблюдаются в структуре гистона H1, причем разные субфракции синтезируются на определенных стадиях клеточного цикла и в определенных тканях

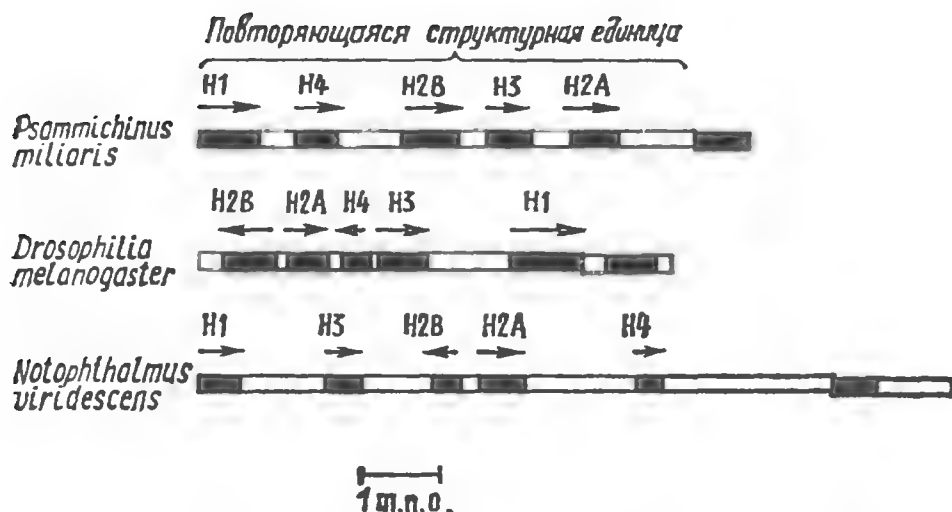


Рис. 122. Схематическое изображение повторяющейся структурной единицы гистоновых генов различных организмов (обратите внимание, что транскрипция разных генов может быть направлена в разные стороны и что порядок расположения генов различен):

сверху вниз: морской еж, плодовая мушка, тритон; черные — транскрибирующиеся области, светлые — спейсеры; стрелки — направление транскрипции

и могут выполнять несколько различную роль. В частности, определенная субфракция гистона H1 синтезируется в интенсивно делящихся клетках млекопитающих и служит маркером злокачественной трансформации.

Синтез гистонов в клетке строго скоординирован с синтезом ДНК: если синтез ДНК подавляется, синтез гистонов падает примерно на 90 %. Остается так называемый базальный уровень синтеза. Возможно, он необходим для восстановления структуры хроматина на репарированной ДНК, для замены деградированных гистонов или для синтеза определенных субфракций. Среди молекул мРНК, кодирующих гистоны, лишь часть несет на 3'-конце поли(А). Возможно, полиаденилирование влияет на время жизни гистоновых матриц и соответственно на уровень и избирательность базального синтеза.

Гистоны, как и другие белки, синтезируются в цитоплазме на рибосомах и в процессе миграции в ядро, посадки на ДНК и даже уже в составе хроматина подвергаются многообразным химическим модификациям. В гистонах происходит ацетилирование лизиновых остатков, снимающее их положительный заряд, метилирование лизиновых, аргининовых и гистидиновых остатков, фосфорилирование треонинов и серинов. К существенному изменению свойств гистонов приводит их поли-ADP-рибозилирование. Иногда к молекулам гистонов в результате модификации присоединяются и более объемистые группы. Например, к боковым цепям некоторых аминокислотных остатков гистонов H2A и H2B присоединяется небольшой белок убихитин длиной в 55 а. о. Модификация ги-

стонов, роль которой пока не установлена, может влиять на их взаимодействие с ДНК, друг с другом, с другими белками и модулировать структуру хроматина. Фосфорилирование гистона H1 строго коррелирует со стадией клеточного цикла: оно происходит за некоторое время до конденсации метафазных хромосом и, видимо, необходимо для этого процесса.

В некоторых дифференцированных клетках, где полностью подавлена транскрипция, встречаются специфические основные белки, частично или полностью заменяющие гистоны. Так, в эритроцитах птиц гистон H1 частично заменен на сходный с ним гистон H5. В сперме рыб гистоны заменены на протамины. Протамины — это белки длиной около 35 а. о., $\frac{1}{3}$ которых представлена основными остатками, главным образом аргинина. Протамины принципиально отличны от гистонов, так как не образуют нуклеосом и плотно упаковывают ДНК иным способом.

ДНК в клетках бактерий также связана с основными белками. Например, в клетках *E. coli* это белки HU1 и HU2. Ничего похожего на нуклеосомы они не образуют, и содержание их в расчете на единицу массы ДНК гораздо меньше, чем гистонов.

2. НЕГИСТОНОВЫЕ БЕЛКИ

Помимо гистонов ядра содержат большое количество разнообразных негистоновых белков. Далеко не все они связаны в ядре с ДНК или с нуклеогистоновой фибриллой. Многие из них входят в состав рибонуклеопротеидов (РНП), которыми богаты транскрипционно активные ядра. Кроме того, в ядре находятся ферменты, непосредственно связанные с функционированием генома: РНК-полимеразы I, II и III; ДНК-полимеразы; топоизомеразы I и II.

Кроме ферментов в ядрах содержатся негистоновые белки, имеющие, по-видимому, отношение к структуре хроматина. К ним относятся так называемые HMG-белки, принадлежащие к двум классам: HMG 14 и 17 и HMG 1 и 2. (Название HMG-белков происходит от англ. high mobility group — группа [белков] с высокой подвижностью, так как в обычных системах гель-электрофореза эти белки движутся быстрее других негистоновых белков хроматина.) Эти белки содержат много положительно и отрицательно заряженных аминокислотных остатков, причем они располагаются асимметрично: N-концевая часть богата кислыми остатками, а C-концевая — основными. Возможно, HMG-белки участвуют в процессах транскрипции и репликации.

3. НУКЛЕОСОМА

Нуклеосомы являются основным регуляторным структурным элементом хроматина. Эти частицы проявляют высокую эволюционную консервативность и обнаруживаются в составе хроматина из лю-

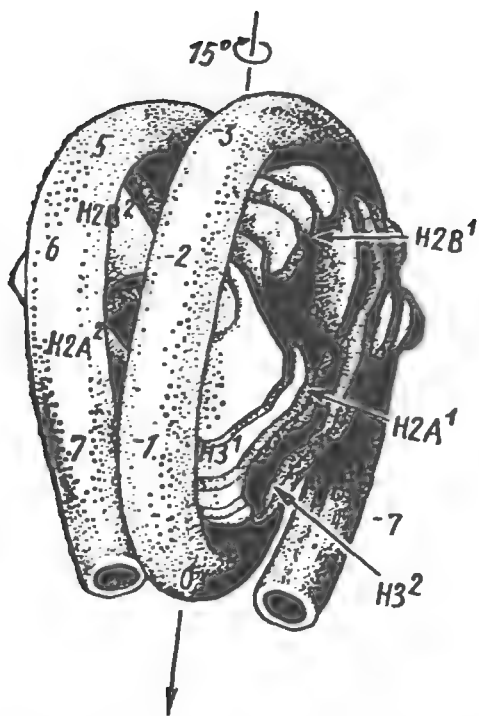


Рис. 123. Пространственная модель нуклеосомы, построенная на основании рентгеноструктурных и некоторых других данных

Модель повернута на 15° относительно оси симметрии, проходящей через середину фрагмента ДНК (на рисунке — сверху вниз); обозначены места контактов молекул гистонов с ДНК: числа от -7 до -7 — полные витки двойной спирали ДНК, между ними по 10 п. о.; за 0 принята середина фрагмента ДНК в нуклеосоме (внизу)

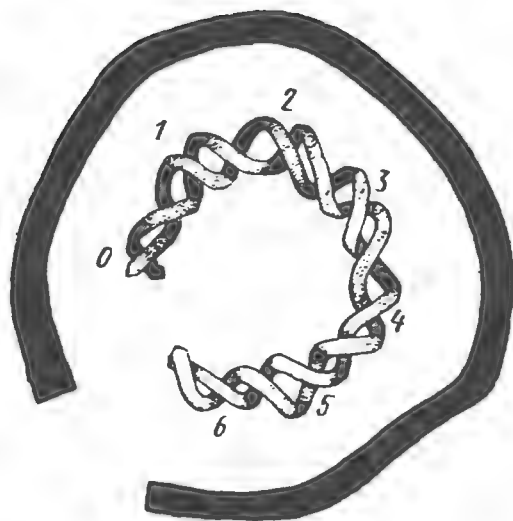


Рис. 124. Ход двойной спирали ДНК в нуклеосоме. (Любезно предоставлено Т. Ричмондом.)

Толстая линия, идущая вокруг двойной спирали, схематически показывает ее форму; хорошо видны «изломы» в участках 1 и 4

бых эукариотических источников. Их можно получить расщеплением хроматина нуклеазами, в частности микрококковой нуклеазой. Минимальная нуклеосома состоит из ДНК длиной 145 п. о. и белковой глобулы. Белковая глобула образована восемью молекулами гистонов: по две молекулы каждого из гистонов H2A, H2B, H3 и H4. Прделанный к настоящему времени в ряде лабораторий рентгеноструктурный анализ кристаллов нуклеосом и их электронная микроскопия пока позволили определить лишь общую форму нуклеосомы (рис. 123). Нуклеосома представляет собой диск диаметром около 11 нм и высотой около 5,7 нм. Гистоны располагаются внутри этого диска, а ДНК навита на поверхность гистонического октамера в виде 1,75 витка левой суперспирали. Вся структура обладает осью симметрии второго порядка. Располагаясь на поверхности, ДНК изгибается неравномерно, и в четырех симметрично расположенных участках наблюдается довольно резкое изгибание ДНК с образованием частичных изломов (рис. 124). Гистоны не полностью заполняют внутреннее пространство нуклеосомы. Белковая глобула имеет рыхлую гидратированную струк-

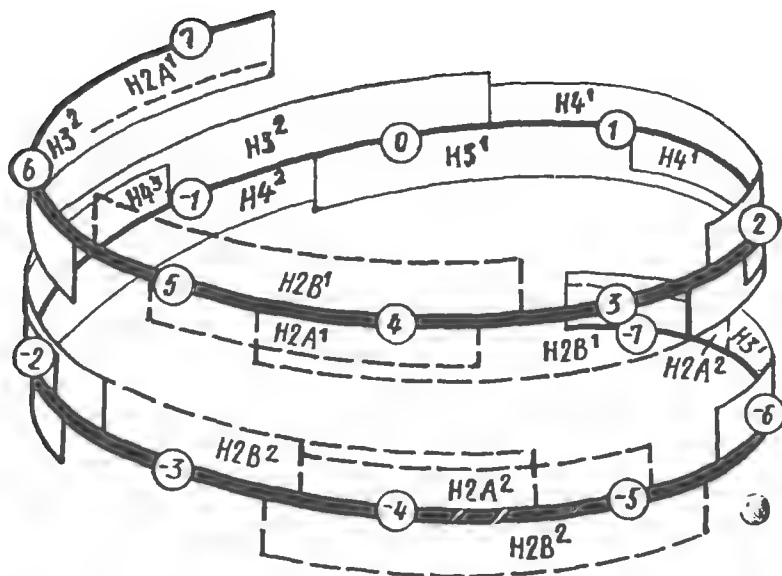


Рис. 125. Карта линейного расположения гистонов на ДНК

[Линейная карта расположения сшивок свернута, как в нуклеосоме. Обратите внимание, что места посадки отдельных молекул гистонов, удаленные друг от друга на линейной карте, сближаются при сворачивании ДНК с шагом 80 п. о. (например, участки контакта гистона H3 с ДНК в середине и на концах нуклеосомной ДНК)]

туру. Пока не удалось получить ее пространственную модель высокого разрешения с помощью рентгеноструктурного анализа; многие важные данные по тонкой структуре нуклеосом были получены с помощью иных подходов.

Один из этих подходов состоит в локализации ковалентных сшивок между нуклеосомными гистонами и ДНК. Принцип локализации заключается в том, что сшитые комплексы белков с ДНК разделяют в двумерных гель-электрофорезах, причем после электрофореза первого направления в геле расщепляют белковый или нуклеиновый компонент комплексов и разделяют во втором направлении только ДНК или только белки соответственно. Таким образом была получена карта линейного расположения гистонов на ДНК (рис. 125). Гистоны H3, H4 располагаются в центре нуклеосомной ДНК, в то время как гистоны H2A, H2B локализованы на периферии. Гистон H3 взаимодействует с центральным и концевым участками нуклеосомной ДНК. Хотя эти участки на развернутой ДНК расположены далеко друг от друга, они сближаются на свернутой в нуклеосому ДНК и, видимо, с ними взаимодействует одна и та же молекула гистона H3. На этой же карте видно, что не вся ДНК сплошь покрыта гистонами, а есть свободные от взаимодействия сегменты, например первые 20 нуклеотидов от 5'-концов обеих цепей нуклеосомной ДНК и участки, расположенные на расстоянии около 120 нуклеотидов от 5'-концов. Внутри нуклеосомы гистоны находятся в тесном контакте друг с другом, о чем свидетельствует образование почти всех возможных

димеров гистонов при белок-белковых сшивках с использованием различных реагентов.

Гистоны проявляют высокую специфичность при взаимодействии друг с другом. При смешивании в растворе наиболее специфичные комплексы возникают при взаимодействии гистонов H3 и H4 с образованием тетрамеров, состоящих из двух молекул каждого из этих гистонов. Гистоны H2A и H2B при взаимодействии образуют высокоспецифичные димеры. При повышении концентрации соли нуклеосомы диссоциируют: сначала происходит отщепление одного димера H2A-H2B, затем второго такого димера и в последнюю очередь диссоциация от ДНК гистонового тетрамера (H3—H4)₂. При понижении ионной силы порядок реассоциации обратный и в конце образуется реконструированная нуклеосома. Реконструкция нуклеосом облегчается в присутствии полианионов, в частности белков, содержащих много сгруппированных в одном месте кислых аминокислот. Реконструкцию нуклеосомы можно проводить не только из ДНК и отдельно взятых димеров и тетрамеров, но также из ДНК и свободных гистонов. Очевидно, структура нуклеосомы в значительной степени определяется гистон-гистоновыми взаимодействиями и структурой гистонового октамера. Так, гистоновый октамер, реконструированный при высокой концентрации соли из гистонов в отсутствие ДНК, по многим свойствам сходен с октамером в составе нуклеосомы. Сборка гистонового октамера происходит за счет взаимодействий центральных гидрофобных сегментов молекул гистонов между собой. Удаление N-концевых участков гистонов с помощью мягкой обработки трипсином не препятствует сборке октамера и даже образованию нуклеосом.

Разработана остроумная генетическая система, позволяющая заменять в клетках дрожжей нормальные гены на их модифицированные аналоги с помощью генно-инженерных манипуляций. В результате в клетке синтезируются измененные белки. Таким образом было показано, что гистоны H2A и H2B дрожжей можно лишить 10—30 концевых аминокислот и что это не влияет на сборку нуклеосом и структуру хроматина и вообще на жизнеспособность клеток. Это особенно странно, если учесть высокую консервативность аминокислотных последовательностей гистонов. Возможно, N-концевые участки нуклеосомных гистонов необходимы не для сборки нуклеосом, а для другой цели, например для транспорта гистонов из цитоплазмы в ядро.

При реконструкции нуклеосом гистоновый октамер может проявлять избирательность к последовательности ДНК. Механизм этой избирательности пока не выяснен в деталях. Предполагается, что физические свойства двойной спирали, в частности ее способность к изгибу (излому), в некоторой степени определяют ее локальную последовательность. Поскольку ДНК навивается на нуклеосому неравномерно, выгодное расположение легко

гнувшихся участков в местах изгибов может увеличить сродство гистонов к данному участку. С другой стороны, не исключено, что октамер просто имеет повышенное сродство к определенным последовательностям ДНК. Например, протяженные полипурин-полипиримидиновые участки длиной свыше 80 п. о. не образуют нуклеосом. Места с правильным чередованием пуринов и пиримидинов, наоборот, по-видимому, благоприятны для образования нуклеосом. Все эти факторы приводят к тому, что посадка нуклеосом на некоторые последовательности ДНК длиной в несколько тысяч пар оснований происходит строго определенным образом, причем регулярность посадки (постоянное расстояние между соседними нуклеосомами) при этом не соблюдается. Это явление называется *фейзингом нуклеосом*, она наблюдается при сборке хроматина *in vivo* и *in vitro*. Однако существуют последовательности ДНК, на которых посадка нуклеосом происходит случайным образом, т. е. фейзинг не наблюдается.

Нуклеосома обладает достаточно высокой стабильностью при различных условиях, однако в ряде случаев были обнаружены сравнительно небольшие конформационные изменения в них. Так, различия в условиях кристаллизации сказываются на взаимодействии одного из гистонов (предположительно H2A) с концевым участком нуклеосомной ДНК. Карта линейной последовательности гистонов на нуклеосомной ДНК также изменяется в деталях в зависимости от того, проводят ли пришивку в ядрах, хроматине (10-нм фибриллах) или выделенных нуклеосомах.

Кроме гистонов в хроматине присутствует большое количество различных негистоновых белков, характер взаимодействия которых с нуклеосомной ДНК пока не ясен. Наиболее богато представлены негистоновые белки HMG 14 и 17, функция которых остается все еще не изученной. HMG 14 и 17 — это близкие по структуре белки, несущие большое количество заряженных групп. Они состоят соответственно из 68 и 74 аминокислотных остатков. Две молекулы этих белков способны к кооперативному связыванию с нуклеосомой, причем каждый белок взаимодействует с концевым участком ДНК и вторым сегментом, расположенным на расстоянии примерно 20 п. о. от ее конца. Эти две области нуклеосомной ДНК в основном свободны от гистонов (см. рис. 125). HMG 14 и 17 связываются с обращенной внутрь нуклеосомы стороной двойной спирали ДНК и не меняют существенным образом общую форму нуклеосомы. Создается впечатление, что эти два белка занимают свободную внутреннюю область ДНК нуклеосомы.

4. ОРГАНИЗАЦИЯ НУКЛЕОСОМНЫХ ФИБРИЛЛ

Нуклеосомы располагаются на ДНК регулярным образом. Разворачивание хроматина при низкой ионной силе позволяет увидеть на электронной микрофотографии нити ДНК, на которых сидят

бусинки нуклеосом. Регулярность расположения нуклеосом на ДНК обнаруживается также при расщеплении хроматина в мягких условиях микрококковой нуклеазой. Если выделить из такого гидролизата ДНК и разделить ее по размерам фрагментов с помощью электрофореза в агарозном геле, то после окрашивания геля видны последовательно расположенные полосы фрагментов ДНК, соответствующих моно-, ди-, три-, тетрануклеосомам и т. д. (рис. 126). Если перенести такие фрагменты на подложку и провести гибридизацию с клонированным фрагментом определенного участка генома, можно получить данные о структуре хроматина в этом участке. В некоторых наиболее упорядоченных участках хроматина лесенка продолжается вплоть до олигомера, содержащего 15 нуклеосом и выше. Однако в большинстве случаев эта лесенка «смазывается» значительно раньше, причем для транскрибируемых участков хроматина регулярность расположения нуклеосом значительно ниже, чем для неактивного хроматина. Длины олионуклеосомных фрагментов ДНК кратны величине, называемой нуклеосомным повтором. Размер нуклеосомного повтора изменяется от 165 п. о. у дрожжей до 200 п. о. у высших эукариот и достигает 240 п. о. в хроматине из спермы морского ежа.

При расщеплении хроматина микрококковой нуклеазой моонуклеосомные частицы вначале содержат фрагмент ДНК, примерно равный по длине нуклеосомному повтору, октамер нуклеосомных гистонов и одну молекулу гистона H1. При более глубоком гидролизе часть ДНК с концов моонуклеосомы расщепляется, теряется гистон H1 и остается так называемая *минимальная нуклеосома*, устойчивая к дальнейшему гидролизу. Минимальная нуклеосома всегда содержит полный октамер гистонов и фрагмент ДНК длиной 146 п. о. независимо от источника хроматина. Таким образом, нуклеосомный повтор складывается из фрагмента ДНК минимальной нуклеосомы строго постоянной длины в 146 п. о. и межнуклеосомной ДНК — так называемого линкера, размер

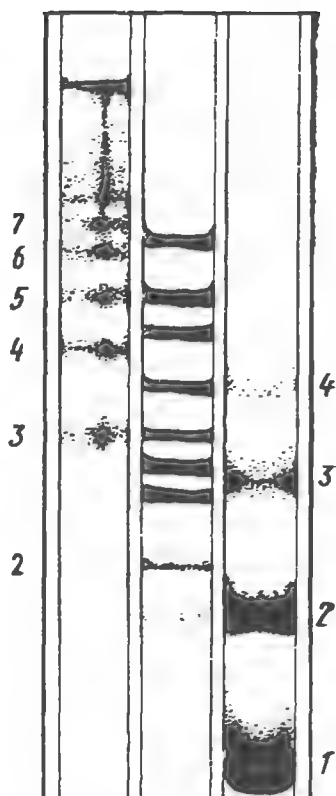


Рис. 126. Электрофорез фрагментов ДНК, образовавшихся при расщеплении микрококковой нуклеазой хроматина печени крысы (левая дорожка) и гриба нейроспоры (правая дорожка)

[В середине нанесены фрагменты ДНК — маркеры молекулярного веса. Этот эксперимент позволяет определить длину нуклеосомного повтора. Для этого определяют длину каждого нуклеосомного фрагмента, делят на его номер (номера указаны рядом с дорожками) и экстраполируют результаты к длинным фрагментам ДНК]

которого зависит от объекта. Он минимален в дрожжах (около 20 п. о.), в клетках млекопитающих обычно равен 50 п. о. и достигает наибольшего размера в хроматине спермы морского ежа (90 п. о.). В действительности длина межнуклеосомной ДНК для каждого объекта является усредненной, а расстояние между соседними нуклеосомами может колебаться вокруг среднего значения, обычно на величину, кратную примерно 10 п. о.

Обработка микрококковой нуклеазой не единственный способ выявить в хроматине регулярное чередование защищенных участков (нуклеосом) и открытых участков (линкеров). Такая структура подтверждается и с помощью некоторых химических проб, которые модифицируют или расщепляют ДНК. Эти соединения расщепляют ДНК там, где она не связана с белками. Гистоны в составе нуклеосомы защищают ДНК, поэтому при ограниченном расщеплении получается характерная нуклеосомная лесенка.

Моно- и олионуклеосомы можно выделить, например, центрифугированием в градиенте сахарозы продуктов расщепления хроматина нуклеазой. Еще более тонкое фракционирование хроматиновых частиц получают с помощью гель-электрофореза, при котором разделяются нуклеосомы с разной длиной линкера или различающиеся по белковому составу.

5. КОНДЕНСАЦИЯ ХРОМАТИНА

При низкой ионной силе (менее 10 мМ NaCl) хроматиновая фибрилла выглядит в электронном микроскопе как бусы нуклеосом, нанизанных на нитку (рис. 127). При повышении ионной силы нуклеосомы сближаются и образуют фибриллы толщиной 10 нм. ДНК в такой фибрилле навита на нуклеосомы в виде более или менее непрерывной спирали с периодом 80 п. о.

В присутствии гистона H1 и ионов Ca^{2+} 10-нм фибрилла образует следующий уровень организации хроматина — фибриллу тол-

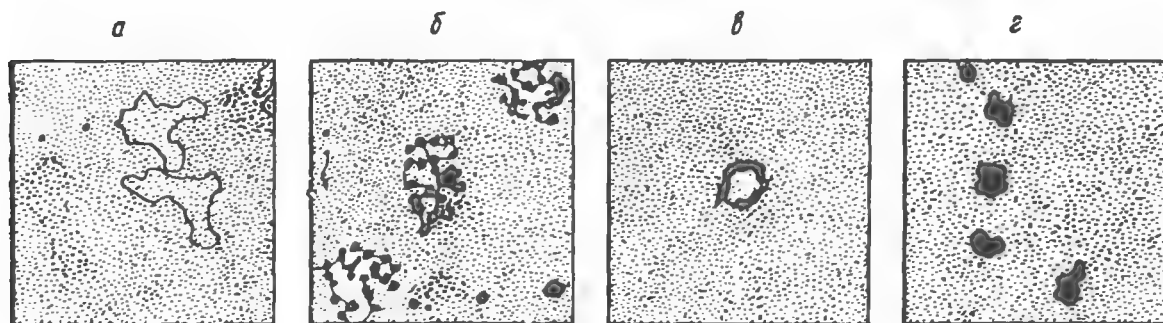


Рис. 127. Простейшая модель хроматиновой фибриллы — так называемая мини-хромосома вируса SV40 (любезно предоставлен С. А. Недоспасовым)

(ДНК этого вируса длиной около 5 т. п. о. реплицируется в ядре клетки-хозяина и представляет собой короткую кольцевую хроматиновую фибриллу): а — ДНК вируса SV40, с которой удалены гистоны; б — мини-хромосомы при низкой ионной силе после удаления гистона H1; в — мини-хромосомы, содержащие H1, при физиологической ионной силе; г — хроматиновые глобулы при той же ионной силе; видны структуры типа «бусы на нитке» (б), фибрилла толщиной 10 нм (г)

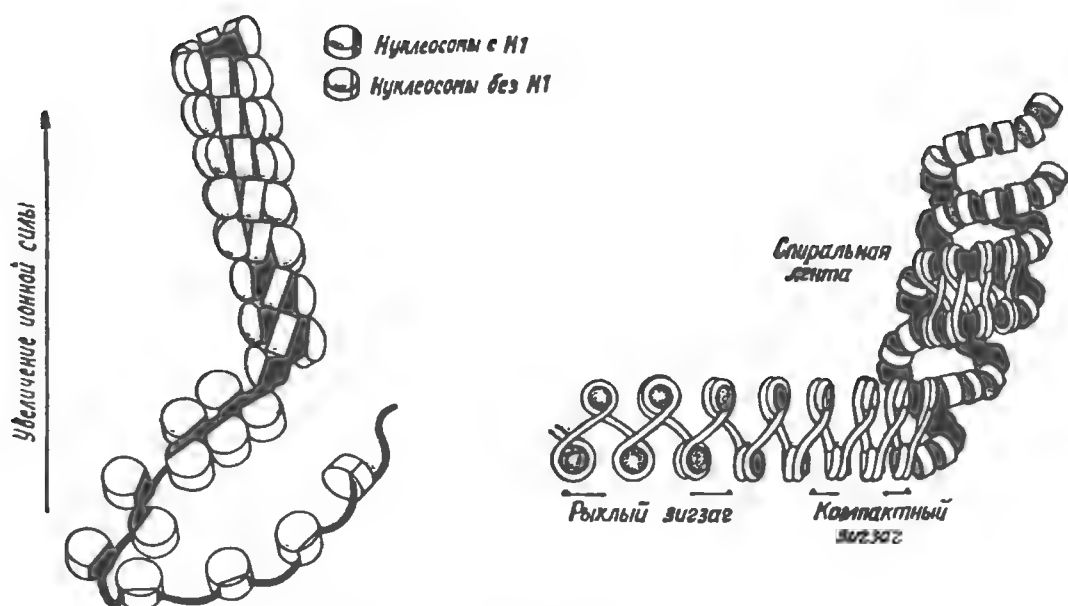


Рис. 128. Модель фибриллы толщиной 30 нм:

а — соленоидная модель; *б* — зигзагообразная модель; на рис. *б* часть нуклеосом с задней стороны спирали не изображена для наглядности.

щиной 30 нм, имеющую, вероятно, форму соленоида. Некоторые исследователи предполагают, что в этом соленоиде группы из нескольких нуклеосом преимущественно взаимодействуют друг с другом, образуя так называемые *нуклеомеры* или *суперглобулы*. Для образования 30-нм фибриллы абсолютно необходим гистон Н1. Глобулярная область гистона Н1 взаимодействует одновременно с обоими концами нуклеосомной ДНК и таким образом фиксирует ее свернутую конформацию в нуклеосоме. Однако для компактизации нуклеосом в составе 30-нм фибриллы необходимы как глобулярная, так и концевые области гистона Н1.

Существует ряд моделей структуры 30-нм фибриллы хроматина. Согласно наиболее обоснованной модели, межнуклеосомная ДНК вместе с нуклеосомной ДНК образуют непрерывную левую суперспираль, в которой соседние нуклеосомы располагаются одна за другой (рис. 128). Согласно другой, «зигзагообразной», модели, межнуклеосомная ДНК образует распрямленные участки, которые связывают соседние нуклеосомы, располагающиеся иным образом, чем в первой модели. В обеих моделях соленоидной структуры на один виток соленоида приходится 6—7 нуклеосом.

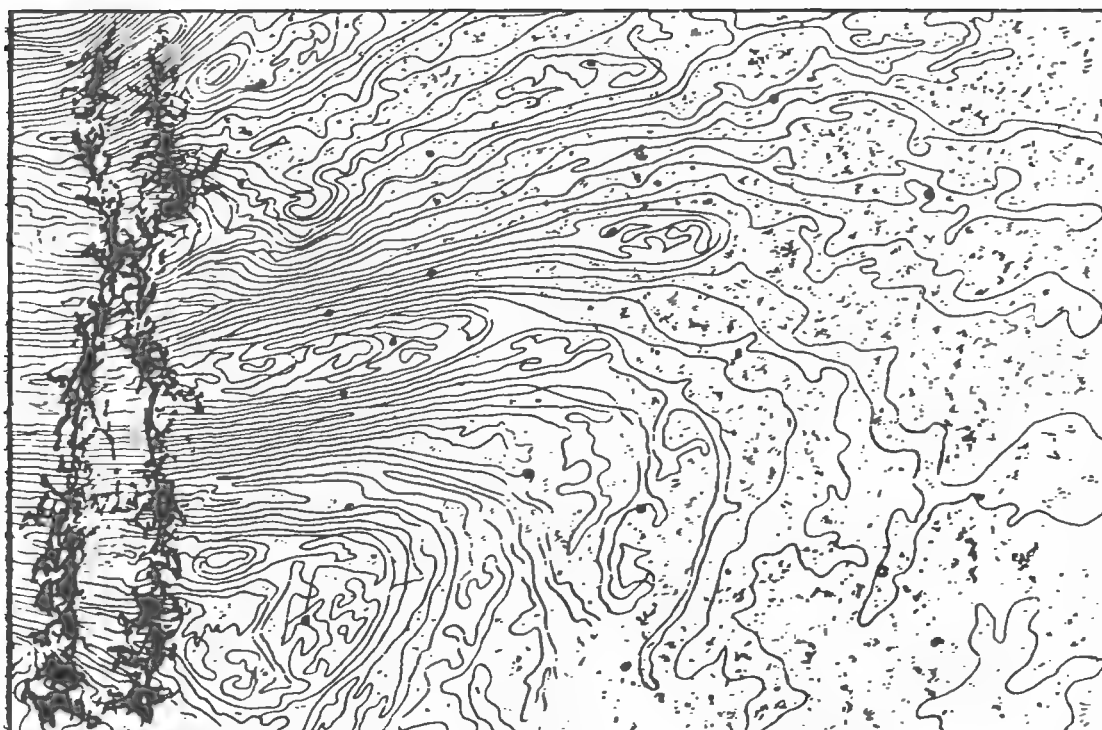
Помимо гистона Н1 в организации соленоидной структуры хроматина участвуют, очевидно, и нуклеосомные гистоны. Положительно заряженные N-концевые области этих гистонов, как упоминалось ранее, несущественны для образования нуклеосомной структуры, но вовлечены в организацию соленоидной структуры хроматина. Удаление этих участков с помощью мягкого расщепления гистонов трипсином в составе хроматина приводит к необратимому разворачиванию соленоида.

6. ДОМЕННАЯ ОРГАНИЗАЦИЯ ХРОМАТИНА

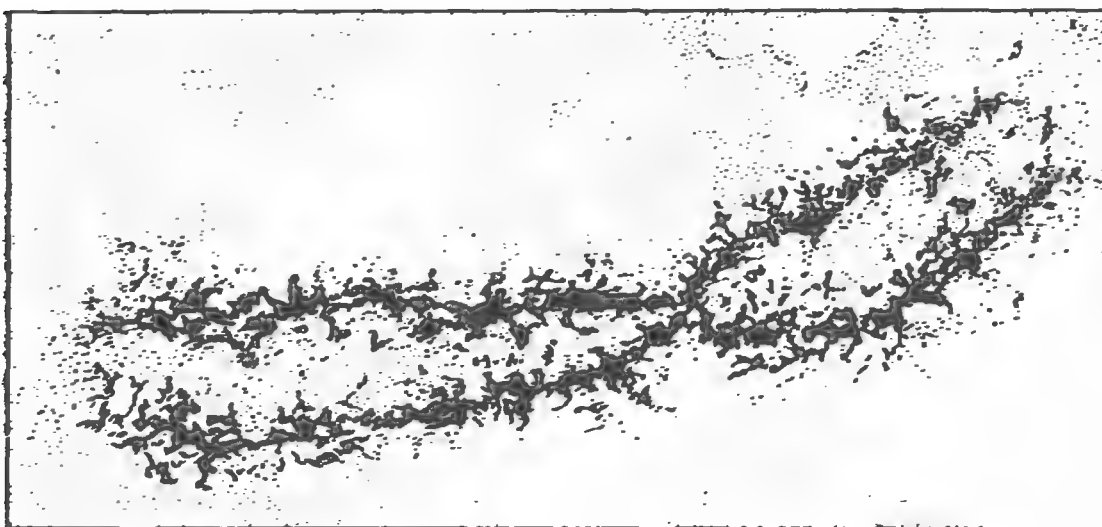
Представление о доменной организации хромосом эукариот было первоначально гипотезой, выдвинутой по аналогии с хорошо установленной доменной структурой бактериального нуклеотида. Хромосома в *E. coli* существует в клетке в виде более или менее компактной структуры. Она состоит из нескольких десятков независимых суперспирализованных петель, которые могут релаксировать по отдельности. Суперспирализованное состояние ДНК, обладающее повышенной энергией, поддерживается в клетках бактерий ферментом ДНК-гиразой, использующим энергию АТФ. В эукариотических клетках этот фермент до сих пор не обнаружен, несмотря на многочисленные попытки его найти.

Домены эукариотической хромосомы отличаются от прокариотических доменов. Представление о доменах прокариотической хромосомы сформулировано на основании опытов по релаксации ДНК. Представление об эукариотических доменах опирается на опыты по электронной микроскопии митотических хромосом, с которых удалены гистоны. ДНК эукариот, точнее нуклеосомная фибрилла, находится в релаксированном состоянии. Обработка релаксирующим ферментом не изменяет ее конформации. Следует учитывать, что ДНК навивается на нуклеосомы спиралью. Если тем или иным способом удалить гистоны с ДНК, то в ней возникают супервитки. Особенно нагляден этот эффект при использовании в качестве модели хроматина кольцевой мини-хромосомы вируса ОВ-40 длиной около 5 т. п. о. Как видно из рис. 127, мини-хромосома на электронных микрофотографиях представляет собой релаксированную структуру. После удаления гистонов ее ДНК суперспирализована. Существует предположение, что транскрипционно активные петли эукариотической хромосомы все-таки находятся в торзионно-напряженном состоянии и релаксируют под действием топоизомераз.

Доменная организация хроматина, как уже говорилось, проявляется на уровне метафазных хромосом. Если обработать метафазную хромосому при высокой концентрации NaCl, удаляющей гистоны, то происходит ее деконденсация и ДНК отделяется от остова метафазной хромосомы в виде большого количества петель (рис. 129). Для сохранения структуры эукариотических хромосом совершенно необходимы ионы Ca^{2+} (и, возможно, Cu^{2+}). Остов сохраняет общую форму метафазной хромосомы и состоит главным образом из полипептидов двух типов. Один из них, по-видимому, топоизомераза II. Если окрасить метафазные хромосомы антителами к топоизомеразе II, выявляется структура, очень сходная с остовом, видимым при удалении гистонов. Начало и конец каждой петли ДНК в таких «распушенных» метафазных хромосомах локализуются в одном и том же месте остова. Размер одной петли для млекопитающих составляет в среднем обычно 40—50 т. п. н.



a



b

Рис. 129. «Распушенная» метафазная хромосома, полученная путем удаления гистонов при высокой концентрации соли (a); такой же препарат, обработанный нуклеазой: белковый остов остается, петли гидролизуются (б)

Возможно, что такой структурный домен является функциональной единицей. Доменная организация ДНК сохраняется, по всей вероятности, также в интерфазном хроматине.

Некоторые биохимические подходы позволяют идентифицировать предполагаемые места прикрепления петель (доменов) к

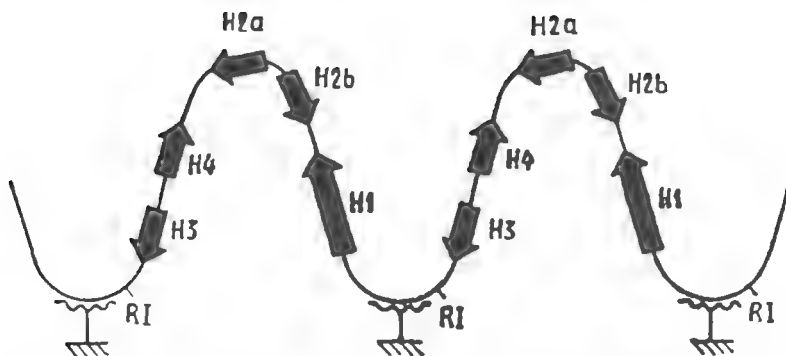


Рис. 130. Предполагаемые места прикрепления гистоновых генов к остоу хромосомы

[Обратите внимание, что они обрамляют всю группу (кластер) генов, а не отдельные транскрипционные единицы.]: R1 — участок расщепления ДНК рестриктазой EcoRI, располагается всегда в одном и том же месте кластера

остову в известных генах. Подобные эксперименты показали, что места прикрепления обрамляют транскрипционные единицы или группы (кластеры) генов (рис. 130).

7. МЕТАФАЗНЫЕ ХРОМОСОМЫ

В конденсированном состоянии каждый домен хроматина представляет собой, вероятно, компактную глобулу, которая занимает в метафазной хромосоме четко определенное положение для каждого участка ДНК. При локализации определенных генов в метафазной хромосоме они всегда обнаруживаются в одном и том же ее участке. Регулярная организация метафазных хромосом подтверждается также тем, что окрашивание их различными красителями дает стандартную картину в виде чередующихся полос более и менее интенсивной окраски. Полученная при окрашивании характерная исчерченность является надежным тестом для идентификации отдельных хромосом.

Благодаря последовательной конденсации хроматина метафазная хромосома, имеющая длину около 5 мкм, содержит ДНК длиной до 5 см, т. е. линейные размеры ДНК уменьшаются примерно в 10 тыс. раз. На рис. 131 иллюстрируются последовательные стадии конденсации.

Механизм конденсации, непосредственно приводящей к образованию метафазных хромосом, пока не выяснен. Очевидно, что существенную роль в этом процессе играет фосфорилирование гистона H1 по сериновым и треониновым остаткам, которое протекает на строго определенной стадии клеточного цикла и предшествует по времени образованию метафазных хромосом.

Как известно, в дрожжах не наблюдается типичная картина митоза и нет метафазных хромосом (возможно, потому что у них очень маленькие хромосомы). По-видимому, с этим связано отсутствие гистона H1 в дрожжах, несмотря на типичную нуклеосомную структуру хроматина.

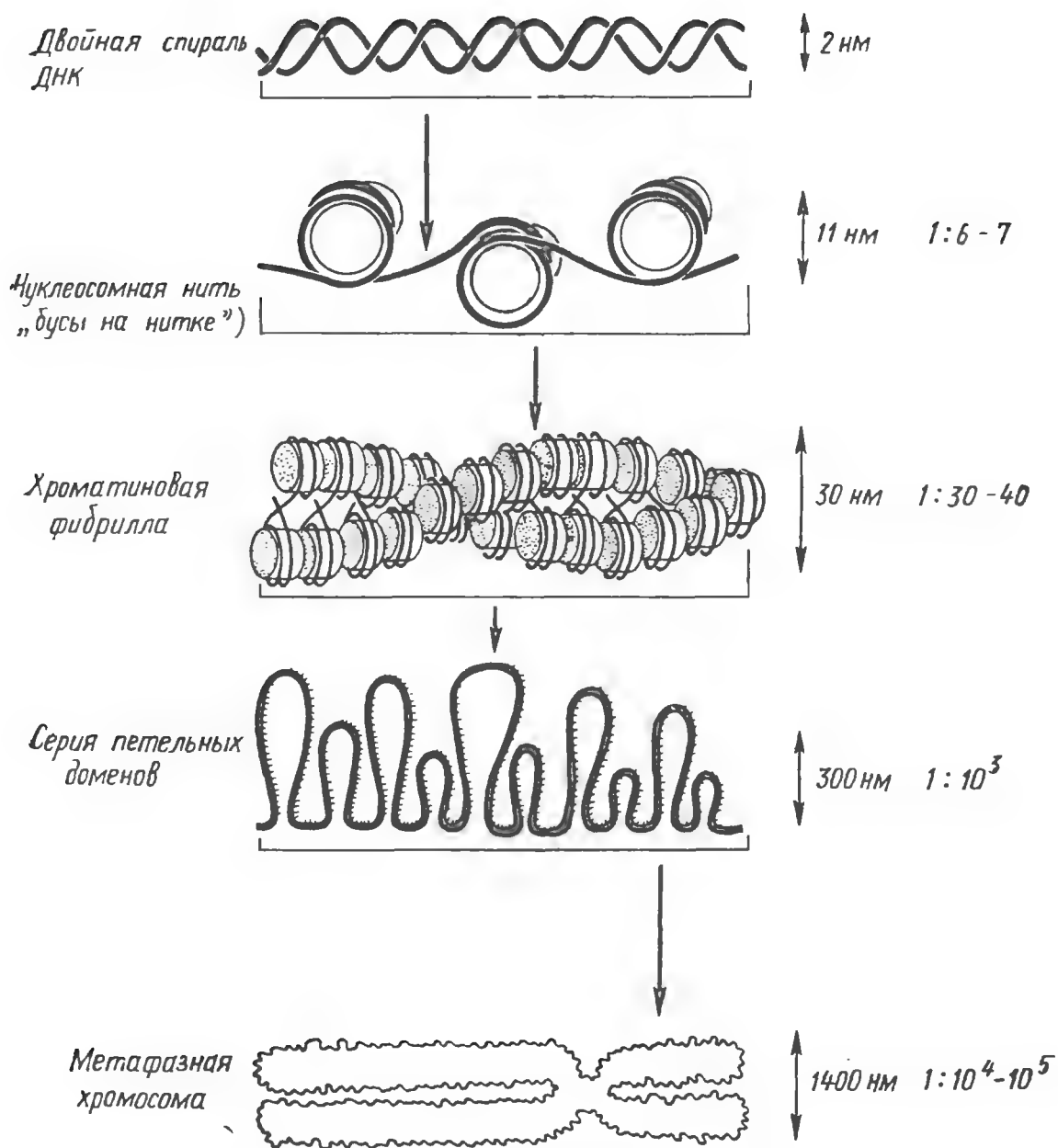


Рис. 131. Уровни компактизации хроматина

Справа — степень компактизации по сравнению с длиной исходной ДНК, которая достигается на каждом этапе

8. РЕГУЛЯТОРНЫЕ БЕЛКИ ХРОМАТИНА

Общие принципы регуляции транскрипции у прокариот хорошо известны (см. гл. 10). Однако многолетние поиски регуляторных белков эукариотической транскрипции пока привели к скромным успехам. У эукариот (за исключением дрожжей) не обнаружено простых регуляторных генетических систем, доступных для биохимического исследования. Поэтому при поиске регуляторных ДНК-связывающих белков эукариот обычно используют различные варианты методов суррогатной генетики.

Некоторые регуляторные «мотивы» (см. гл. X, раздел 1) встре-

чаются в количестве десятков тысяч и более на гаплоидный геном, и число молекул белков, узнающих эти «мотивы», соответственно велико. Для выделения этих белков используют различные методы обнаружения их специфических комплексов с небольшими мечеными фрагментами ДНК, содержащими «мотивы». Второй подход — аффинная хроматография экстрактов ядер, содержащих негистоновые белки, на колонках с пришитыми короткими синтетическими или природными фрагментами ДНК. Эти фрагменты содержат последовательность узнавания. Третий подход — различные методы ковалентной сшивки белков с ДНК *in vivo* и *in vitro* в месте их связывания с последующей идентификацией белкового компонента. Одновременно, как правило, исследуют влияние изучаемых белков на транскрипцию в системе *in vitro*. Таким образом удалось очистить так называемый белок Sp1, который стимулирует транскрипцию с одного из сильных промоторов животных клеток — раннего промотора вируса ОВ-40, а также, по-видимому, с ряда других клеточных промоторов. Он узнает в них мотив GGGCGG.

В то же время многие регуляторные белки эукариот, как и у прокариот, составляют ничтожную долю всех белков. Регуляторные последовательности эукариотических генов иногда удалены от промотора на значительное расстояние (энхансеры, см. гл. X, раздел 2) и расположены впереди или позади него. Это затрудняет поиск белков, узнающих определенные последовательности ДНК. ДНК в хроматине свернута в спираль с шагом около 80 п. о., и сайт узнавания может быть составлен из фрагментов разных участков, разнесенных в первичной структуре на эту величину. Поэтому изучение и моделирование механизма узнавания ДНК в хроматине требуют разработки совершенно новых подходов.

Пожалуй, наиболее изученный регуляторный белок эукариот — фактор транскрипции генов 5S РНК шпорцевой лягушки. Его структура и механизм действия рассмотрены в гл. X.

Еще одна регуляторная система, механизм которой проясняется, — это гены, стимулируемые стероидными гормонами. Белок-регулятор, связываясь с гормоном, приобретает способность садиться на регуляторные участки и активировать соответствующие гены. Описан также белок, который связывается с регуляторной областью генов теплового шока дрозофилы и включает эти гены (см. с. 199). Тепловой шок вызывает переход этого белка из цитоплазмы в ядро и его связывание с последовательностью ДНК, обеспечивающей включение генов теплового шока. Предполагается, что, связываясь с ДНК вблизи промотора, он активирует РНК-полимеразу II.

В дрожжах обнаружены генетические регуляторные элементы типа прокариотических репрессоров. Некоторые из них интенсивно изучаются, в частности репрессор гена GAL. Этот белок состоит по крайней мере из двух функциональных доменов. Один

из них отвечает за транспорт белка из цитоплазмы в ядро, другой — за связывание с ДНК.

Гены, кодирующие рибосомные РНК, в той или иной степени всегда активны. Как правило, их регуляция осуществляется путем изменения числа работающих генов, а не интенсивности их транскрипции. Организация рибосомных генов и расположение регуляторных последовательностей были рассмотрены ранее (см. гл. X, раздел 3). Структура хроматина рибосомных генов обладает несколькими важными особенностями. Во-первых, активные рибосомные гены практически лишены гистонов и покрыты работающими молекулами РНК-полимеразы I, идущими одна за другой. Во-вторых, промотор рибосомных генов, а также его гомологи в спейсере всегда свободны от гистонов и находятся в активной конформации. В-третьих, существование дополнительных мест инициации и особенно терминации транскрипции в спейсере необходимо для нормальной работы промотора. Этот факт не совсем понятен. Возможно, функциональные участки спейсера служат местами сборки комплекса преинициации РНК-полимеразы I и необходимых ей факторов или просто повышают локальную концентрацию фермента; возможно также, что они поддерживают ДНК в активном, свободном от гистонов состоянии и тем самым способствуют быстрому началу транскрипции в нужный момент. Существует гипотеза, что близлежащие дополнительные промоторы могут активировать и другие РНК-полимеразы. Быстрый прогресс в изучении GAL-репрессора позволил показать, что этот репрессор работает в чужеродных животных и растительных клетках, если в них введены соответствующие регуляторные последовательности.

9. СТРУКТУРА АКТИВНОГО ХРОМАТИНА

Несколько последовательных уровней конденсации хроматина переводят его в высококомпактное состояние. Конденсация, очевидно, препятствует доступу к активным участкам хроматина молекул РНК-полимераз и их продвижению вдоль ДНК, сопровождающему транскрипцию, поэтому транскрипция может проходить лишь на деконденсированных участках хроматина. Таким образом, деконденсация может быть важным уровнем сложной иерархии процессов регуляции активности эукариотических генов. Наиболее наглядно деконденсация хроматина при интенсивной транскрипции наблюдается электронно-микроскопически на рибосомных генах, активно транскрибирующихся в осцитах шпорцевой лягушки *Xenopus* (рис. 132). Эти гены плотно заполнены молекулами РНК-полимеразы, с которых свисают вновь синтезированные молекулы РНК в комплексе с белками, так что общая картина этих транскрибированных генов напоминает новогоднюю елочку. Длина транскрибируемого участка почти достигает длины соответствующей ДНК, т. е. ДНК практически не конденсирована, и между молеку-

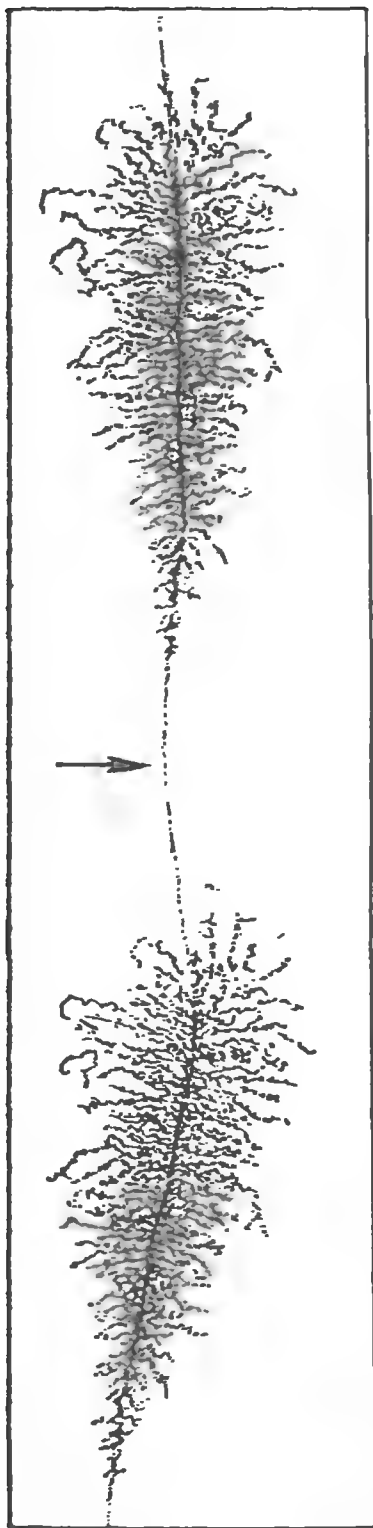


Рис. 132. Интенсивная транскрипция рибосомных генов
Видны градиент длины РНП вдоль гена и нетранскрибируемый спейсер (отмечен стрелкой)

лами РНК-полимеразы не видно нуклеосом, а также отсутствуют нуклеосомы в какой-либо иной конформации. Аналогичную картину можно наблюдать на электронных микрофотографиях генов, транскрибируемых РНК-полимеразой II, например в петлях хромосом типа ламповых щеток (рис. 133). В этих хромосомах активные домены образуют петли длиной в десятки тысяч пар нуклеотидов, на которых можно наблюдать процесс транскрипции.

Деконденсацию хроматина при транскрипции можно также наблюдать с помощью светового микроскопа на политенных хромосомах дрозофилы. Такие хромосомы содержатся во многих тканях личинок насекомых. Политенные хромосомы дрозофилы состоят примерно из 1000 нитей ДНК, лежащих рядом друг с другом таким образом, что гомологичные участки соседствуют и образуют поперечные полосы. Политенные хромосомы соответствуют интерфазному хроматину. Каждый функциональный домен в политенной хромосоме представлен в виде диска, содержащего плотноупакованную ДНК. Диски разделены менее плотными междисковыми участками. Чередование дисков и междисков образует характерную строго постоянную картину, причем крупные генетические перестройки проявляются в видимых изменениях хромосом. В ходе индивидуального развития личинок картина дисков и междисков несколько меняется. Но особенно ярко изменения транскрипционной активности хроматина политенных хромосом проявляются

при индукции генов. Такая индукция достигается, например, при нагревании личинок (так называемый тепловой шок) или при введении гормона насекомых экдизона. При активации транскрипции происходит резкая деконденсация хроматина в определенных дисках и образуются так называемые пухы. В пухах можно обна-

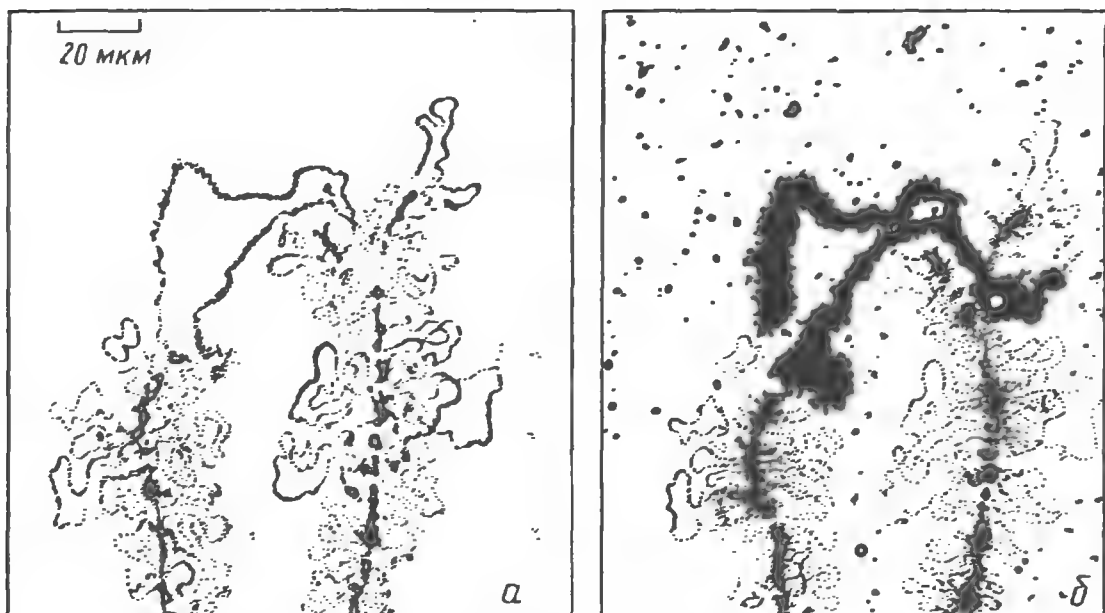


Рис. 133. Хромосома типа ламповой щетки (а); транскрипция гистоновых генов на одной из петель такой хромосомы (б)

От осевой структуры (рис. а) отходят многочисленные петли активно транскрибируемого хроматина; зернами радиоактивной метки помечена новообразованная гистоновая мРНК (рис. б)

ружить с помощью различных молекулярно-биологических методов молекулы вновь синтезированной РНК, связанной со специфическими белками, а также РНК-полимеразы II и топоизомеразы I. Связанная с транскрипцией деконденсация хроматина является обратимым процессом. Это легко прослеживается на генах дрозофилы, которые активируются при тепловом шоке: при понижении температуры с 37 °С до нормальной для дрозофилы температуры 24 °С наблюдается регрессия пухов в конденсированное состояние, происходящая в течение нескольких часов.

Активные в транскрипции участки хроматина отличаются от неактивных участков по многим свойствам. Это может являться следствием деконденсации, присутствия ряда новых белков или изменения состояния ДНК. Активный хроматин существенно быстрее гидролизруется нуклеазами, такими, как ДНКаза I и микрококковая нуклеаза. В неактивном хроматине не обнаруживаются молекулы РНК-полимеразы. Однако при активации наблюдается появление на гене молекул РНК-полимеразы II, причем их количество пропорционально интенсивности транскрипции. В слаботранскрибирующихся генах может присутствовать одна или несколько молекул, в то время как в активно транскрибирующихся генах рРНК или в генах, активированных при тепловом шоке, ДНК буквально насыщена РНК-полимеразами: на участок гена длиной в 60—100 п. о. в среднем приходится по одной молекуле РНК-полимеразы. Как уже говорилось, с активным хроматином ассоциированы также молекулы топоизомеразы I. При активации генов в регуляторных участках хроматина (в области

промоторов и энхансеров) появляются помимо молекул РНК-полимераз также некоторые специфические регуляторные белки. Их появление координирует с возникновением так называемой гиперчувствительности — резким возрастанием чувствительности некоторых участков в регуляторных областях к действию ДНКаз.

Активация хроматина сопровождается также локальным ацетилированием N-концевых областей гистонов H2A, H2B, H3 и H4. Вероятно, активный и неактивный хроматин различаются и по содержанию гистоновых вариантов. Например, в полностью репрессированном хроматине эритроцитов цыпленка гистон H1 частично заменен близким ему по структуре гистоном H5.

Имеются различия и в состоянии ДНК транскрибируемых и не-транскрибируемых участков генома. Так, неактивная в транскрипции ДНК в большей мере метилирована, в то время как содержание 5-метилцитозина существенно меньше в транскрибируемых генах и особенно в их регуляторных последовательностях. При действии топоизомеразы I на неактивный хроматин количество супервитков не меняется, что свидетельствует о том, что эта ДНК находится в релаксированном состоянии, в ней нет торзионных напряжений, которые бы снимались под действием топоизомеразы I. Возможно, что в активном хроматине возникают такие торзионные напряжения, которые облегчают связывание регуляторных белков и работу РНК-полимеразы.

10. СУДЬБА НУКЛЕОСОМ ПРИ ТРАНСКРИПЦИИ

Важный вопрос организации хроматина касается судьбы нуклеосом при транскрипции. Электронная микроскопия интенсивно транскрибирующихся участков хроматина, например рибосомных генов, ясно показывает, что нуклеосом на них нет даже в тех случаях, когда между молекулами РНК-полимеразы, движущимися одна за другой по гену, виден промежуток. Необходимо отметить, что регуляция активности рибосомных генов осуществляется в клетке путем изменения числа работающих генов, но не интенсивности транскрипции. Однако промоторы рибосомных генов всегда находятся в активной конформации (свободны от гистонов).

Нуклеосомы отсутствуют и в области потенциально активного промотора РНК-полимеразы II, даже когда транскрипция еще не началась. Так организована область вблизи начала транскрипции ранних генов SV 40 (она же является областью начала репликации). Не обнаруживаются нуклеосомы и в регуляторных участках генов овальбумина кур, активных в клетках яйцевода, а также глобиновых генов.

Прямым методом изучения наличия гистонов в тех или иных участках ДНК является ковалентная сшивка гистонов с ДНК.

При подробном изучении этим методом генов теплового шока оказалось, что промоторная область всегда преимущественно свободна от гистонов, а участок транскрипции неиндуцированного гена находится в обычной нуклеосомной конформации. При индукции гистоны сбрасываются с ДНК, а ген покрывается молекулами РНК-полимеразы, идущими одна за другой. Гистоны удаляются с гена в определенном порядке: прежде всего теряется гистон H1, чувствительный к разрушению 30-нм фибриллы, затем, по-видимому, удаляются H2A и H2B и последними снимаются H3 и H4.

Однако полное удаление гистонов имеет место лишь в немногих случаях при максимальной интенсивности транскрипции. Как показали многочисленные эксперименты, при умеренной и слабой транскрипции нуклеосомы (гистоны) сохраняются на ДНК. Это подтверждают и биохимические данные, и электронная микроскопия, причем структура этих нуклеосом, вероятно, не отличается от обычных нуклеосом неактивного хроматина.

Представленные выше данные позволяют сформулировать следующую модель структурных переходов в хроматине при транскрипции. Два крайних состояния представлены неактивным в транскрипции хроматином, когда он организован в виде 30-нм фибрилл, упакованных каким-то образом на следующем уровне организации хроматина. Другое крайнее состояние хроматина соответствует полностью линеаризованной ДНК максимально активных генов, таких, как рибосомные гены ооцитов *Xenopus* или полностью активированные гены теплового шока дрозофилы. Эти гены заполнены плотно сидящими друг за другом молекулами РНК-полимераз, и свободные от полимераз участки настолько малы, что их недостаточно для образования нуклеосом или хотя бы связывания гистонов. В случае умеренной транскрипции, когда количество молекул РНК-полимераз на ген невелико, размеры свободной от полимераз ДНК достаточно велики, чтобы на них собирались нуклеосомы. Однако такие одиночные нуклеосомы или небольшие группы нуклеосом не содержат гистона H1 и компактизованы только в 10-нм фибриллы. Гистон H1 взаимодействует с 10-нм фибриллой, по-видимому, кооперативно: для его связывания требуется участок, содержащий по крайней мере 6—7 нуклеосом подряд, что достаточно для образования одного витка соленоида в 30-нм фибрилле хроматина. Такие компактные 30-нм фибриллы могут образовываться только на участках хроматина, транскрибируемых с очень низкой эффективностью, соответствующей, вероятно, присутствию одной молекулы РНК-полимеразы на 2 тыс. п. о. и более транскрибируемой ДНК. Согласно этому представлению, РНК-полимераза временно смещает гистоны с занимаемых ими участков на какие-то промежуточные отрицательно заряженные акцепторы. Ими могут быть любые облегчающие реконструкцию хроматина полианионные молекулы, такие, как РНК или кислые белки типа HMG или нуклеопазмина (см. раздел 12 этой главы).

Эта модель структурной динамики транскрипционно активного хроматина не является единственной. Так, в активно транскрибируемом хроматине рибосомных генов гриба *Physarum* обнаружены развернутые нуклеосомы, в которых гистоны остаются связанными в частично или полностью линейаризованной ДНК нуклеосомы. Эта модель предполагает, что в процессе транскрипции происходит линейаризация ДНК, но РНК-полимераза не смещает молекулы гистонов с транскрибируемых участков. Напомним, что регуляторный белок TFIIIA генов 5S РНК шпорцевой лягушки прочно связывается с регуляторным участком, лежащим в транскрибируемой области, и не диссоциирует при прохождении РНК-полимеразы III.

Влияние степени конденсации хроматина на транскрипцию ярко проявляется в так называемом эффекте положения. Эффект положения открыт и лучше всего изучен у дрозофилы. В хроматине дрозофилы имеются участки сильно конденсированного хроматина, который не транскрибируется,— так называемого *гетерохроматина*. С помощью различных хромосомных перестроек можно переместить тот или иной участок активного в транскрипции хроматина и поместить его рядом с гетерохроматином. Это приводит к подавлению транскрипции активного участка в большей или меньшей степени в зависимости от расстояния до гетерохроматинового участка: чем он ближе, тем эффективнее подавляет транскрипцию.

Эффект положения проявляется и на уровне отдельных генов. Так, часть рибосомных генов дрозофилы содержит небольшие гетерохроматиновые вставки. Конденсированное, неактивное состояние этих вставок распространяется на значительное расстояние вдоль хроматиновой фибриллы и подавляет транскрипцию на расстоянии нескольких тысяч пар нуклеотидов.

11. ГИПЕРЧУВСТВИТЕЛЬНЫЕ МЕСТА ХРОМАТИНА

Существование гиперчувствительных к нуклеазам свободных от нуклеосомных гистонов участков в промоторной и регуляторной областях — обязательный признак активного или потенциально активного в данных клетках гена. Отсутствие гистонов в этих участках может быть обусловлено различными механизмами. Во-первых, некоторые специфические последовательности ДНК, такие, как полиdA: поли dT, не способны в силу некоторых причин (например, излишней жесткости) сворачиваться в нуклеосомную структуру. Такие свободные от гистонов области гомополимеров найдены, в частности, в промоторных областях некоторых дрожжевых генов, и их присутствие существенно для активного состояния этих генов. Другим механизмом может быть связывание специфических белков с регуляторными участками ДНК, что препятст-

вует посадке нуклеосом. Наконец, третья возможность предусматривает фиксированное расположение нуклеосом на ДНК таким образом, что между некоторыми нуклеосомами имеются участки межнуклеосомной ДНК, недостаточные для связывания еще одной нуклеосомы. Например, нуклеосомы садятся на расположенные тандемно множественные гены 5S РНК в строго определенных местах, даже при реконструкции на «голой» ДНК. При этом промоторная область гена всегда оказывается в межнуклеосомном промежутке и доступна для связывания с регуляторным белком TFIIIA, необходимым для активации гена 5S РНК.

Возникновение гиперчувствительных мест в хроматине может быть обусловлено не только отсутствием нуклеосом, но и другими факторами: особой локальной конформацией ДНК и/или связыванием негистоновых белков. Хотя большая часть ДНК в хроматине находится в обычной В-конформации, некоторые ее участки могут принимать другую конформацию, например из-за особенностей последовательности. Так, протяженные (более 8—10 п. н.) участки строгого чередования пуринов и пиримидинов в экспериментах *in vitro* при повышении ионной силы склонны переходить в левоспиральную Z-конфигурацию. Еще одна необычная конформация ДНК — так называемая Н-форма (см. гл. 1). В→Н-переход наблюдается в полипурин-полипиримидиновых последовательностях при понижении рН. Однако до сих пор не ясно, возможны ли такие переходы в клетке. Известно, что многие нуклеазы с высокой избирательностью расщепляют ДНК на границе между участками с различной конформацией, например В—Z или В—Н.

Регуляторные ДНК-связывающие белки прокариот вызывают заметные изменения конформации ДНК. При рентгеноструктурном исследовании комплекса регуляторного белка TFIIIA со своим участком ДНК оказалось, что двойная спираль находится в А-форме. Другие изменения (изгибы и изломы двойной спирали) можно обнаружить с помощью электронной микроскопии, электрофореза ДНК-белковых комплексов, а также при действии нуклеаз. Связанный белок защищает от расщепления 15—30 п. о. в месте связывания и порождает два участка повышенной чувствительности к нуклеазам с обеих сторон от места связывания. Тонкий анализ мест гиперчувствительности в хроматине эукариот показал, что они имеют точно такую же структуру — две гиперчувствительные точки, разделенные защищенным участком.

12. РЕПЛИКАЦИЯ ХРОМАТИНА

Непосредственно на участке репликационной вилки нуклеосом не видно, что связано, скорее всего, с вытеснением гистонов с участка репликации. Нуклеосомы появляются на новообразованных нитях ДНК в минихромосоме SV 40 на расстоянии примерно 200—400 п. о. от репликационной вилки. Судьба старых нуклеосом

в хроматине тоже достаточно хорошо известна. В процессе репликации не происходит смешивания старых и новых гистонов в составе октамеров. Такая консервативность старых нуклеосом может быть прослежена в течение нескольких клеточных генераций без их разбавления вновь синтезированными гистонами. Однако распределение имеющихся гистоновых октамеров по двум дочерним нитям ДНК происходит, скорее всего, случайным образом.

Для правильной посадки гистонов на ДНК *in vitro* необходимы промежуточные полианионные акцепторы. В их присутствии гистоны могут связываться с ДНК, давая нуклеосомный повтор, свойственный данному хроматину *in vivo*. В клетке для этого, по-видимому, служат специальные кислые белки. Такой белок, *нуклеоплазмин*, был обнаружен в ооцитах шпорцевой лягушки (*Xenopus*). Он, вероятно, участвует также в транспорте гистонов из цитоплазмы в ядро.

С механизмом клеточной дифференцировки связан интересный вопрос: сохраняется ли на уровне структуры хроматина память об активном или неактивном состоянии гена при клеточном делении и транскрипции? При клеточном делении хроматин, видимо, сохраняет особенности своей структуры, например гиперчувствительные участки в хроматине некоторых генов сохраняются в метафазных хромосомах в тех же местах, что и в интерфазном хроматине. Очевидно, это определяется тем, что регуляторные белки, связанные с промоторными участками генов, ассоциированы с ДНК и в составе метафазной хромосомы. Однако судьба регуляторных белков в процессе репликации ДНК неизвестна.

Модель прямого запоминания в процессе репликации должна предусматривать сохранение регуляторных белков и гистонов в их модифицированном состоянии на активных участках хроматина. Однако пока отсутствуют прямые экспериментальные данные, подтверждающие это предположение.

Еще один возможный механизм сохранения информации об активности генов в ходе клеточного деления — это метилирование ДНК. У прокариот метилаза узнает «полуметилированный» по одной цепи ДНК сайт после репликации и восстанавливает общую картину метилирования. Возможно, сходные механизмы действуют у эукариот. Ряд данных указывают на то, что ингибиторы метилирования ДНК активируют многие гены после одного или нескольких раундов репликации. В растительных клетках метилирование регуляторных участков некоторых генов приводит к их полному выключению на протяжении многих поколений. Это явление трудно отличить от истинной мутации.

ЛИТЕРАТУРА

Албертс Б., Брей Д., Льюис Дж. и др. Молекулярная биология клетки.— М.: Мир. 1986. Т. 2.

Базыкин С. Г. Динамическая структура хроматина // Итоги науки и техники. Сер. Молекулярная биология.— М.: ВИНТИ, 1988. Т. 26. С. 5—121.

Бакаев В. В. Исследование структурно-функциональных отношений в хроматине эукариот // Итоги науки и техники. Сер. Биологическая химия.— М.: ВИНТИ, 1982. Т. 16. С. 128—170.

Недоспасов С. А. Минихромосома как структурно-функциональная модель клеточного хроматина // Итоги науки и техники. Сер. биол. химия.— М.: ВИНТИ, 1982. С. 171—210.

Преображенская О. В., Карнов В. Л., Мирзабеков А. Д. Структура хроматина // Ж. Всесоюзного химич. общ. им. Д. И. Менделеева. 1984. Т. 29. Вып. 2. С. 171—175.

McGhee J. D., Felsenfeld G. Nucleosome structure // Ann. Rev. Biochem. 1980. 49. P. 1115—1156.

Igo-Kemene T., Hörz W., Zachau H. G. Chromatin // Ann. Rev. Biochem. 1982. 51. P. 89—121.

Eissenberg J. C., Cartwright I. L., Thomas G. H., Elgin S. C. R. Selected topics in chromatin structure // Ann. Rev. Genet. 1985. 19. P. 485—536.

Pederson D. S., Thoma F., Simpson R. T. Core particle, fiber, and transcriptionally active chromatin structure // Ann. Rev. Cell Biology. 1986. 2. P. 117—147.

Часть

В

Биосинтез вирусных нуклеиновых кислот

Глава XIII

РЕПЛИКАЦИЯ И ТРАНСКРИПЦИЯ ВИРУСНЫХ ГЕНОМОВ

Для понимания принципов биосинтеза нуклеиновых кислот важно представлять, что в природе существует большое разнообразие способов образования молекул ДНК и РНК. Это разнообразие наиболее подробно изучено в случае репликации и транскрипции вирусных геномов.

1. РЕПЛИКАЦИЯ ГЕНОМА ДНК-СОДЕРЖАЩИХ ВИРУСОВ (ДНК→ДНК)

В этом разделе рассмотрены способы репликации вирусных ДНК-геномов, при которых синтез ДНК происходит исключительно на ДНК-матрице. Репликация вирусных ДНК, протекающая с участием механизмов обратной транскрипции, будет описана ниже (см. с. 308).

Виды вирусных ДНК-геномов

У разных ДНК-содержащих вирусов геном может быть представлен молекулами разных типов. Так, вирионная ДНК паповавирусов [например, обезьяньего вируса 40 (SV40)] представляет собой *кольцевую двухнитевую* молекулу, содержащую несколько более 5 т. п. н. Обе цепи, входящие в кольцевой дуплекс, ковалентно-непрерывны.

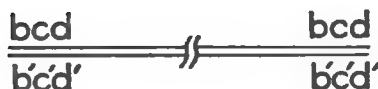
Наиболее известные представители вирусов с *кольцевым однонитевым* ДНК-геномом — мелкие фаги ϕ X174, M13 и др. В состав молекулы ДНК этих фагов входит ~5,4 т. н. Подобную форму имеют и геномы некоторых вирусов эукариот, например гемини-вирусов (вирус золотистой мозаики фасоли и др.).

Весьма широко среди вирусов распространены *линейные двухнитевые* ДНК-геномы. Их размеры колеблются в широких пределах, чаще всего от 30 до 150 т. п. н. Из геномов с полностью из-

1



2



3



4



5



Рис. 134. Концевые участки вирусных линейных двухнитевых ДНК-геномов: 1 — прямой концевой повтор (например, у ДНК фага Т7); 2 — популяция молекул с кольцевыми перестановками и прямыми концевыми повторами (ДНК фага Т4); 3 — «липкие» концы (ДНК фага λ); 4 — инвертированный концевой повтор и ковалентносвязанный терминальный белок (ДНК аденовируса); 5 — ковалентная связь между концами комплементарных цепей (ДНК вируса осповакцины)

вестной первичной структурой самый крупный принадлежит вирусу Эпштейна — Барр — более 172 т. п. н.

Как будет видно из дальнейшего, особое значение для механизмов репликации линейных молекул ДНК имеет структура их концевых участков. У линейных ДНК-геномов не бывает «невыразительных» концов. Соответствующие участки (рис. 134) могут иметь *прямые концевые повторы* длиной от сотни и более (например, ДНК фага Т7) до тысяч (Т-четные фаги и др.) пар нуклеотидов. При этом если у фага Т7 все геномные молекулы ДНК идентичны, то молекулы ДНК Т-четных фагов существенно различны, даже когда они образованы в одной клетке, зараженной единственной фаговой частицей; геномы Т-четных фагов (и ряда других вирусов) характеризуются так называемыми *кольцевыми перестановками*. Еще один вариант концевой структуры вирионных ДНК-дуплексов — «липкие» (т. е. взаимно комплементарные) *однонитевые концы*, длина которых обычно находится между 10 и 20 нуклеотидами (фаги λ , Р2, Р4), но может укорачиваться до одного нуклеотида (герпес-вирусы), если в этом случае вообще позволительно называть такие концы «липкими».

В ДНК некоторых вирусов имеются *инвертированные концевые повторы*, длиной от нескольких нуклеотидов (фаг ϕ 29) до 100—150 нуклеотидов (аденовирусы) и больше; в обоих указанных случаях

геномы имеют важную особенность: к 5'-концам цепей ДНК ковалентно присоединены вирус-специфические белки. Весьма своеобразны концы генома вируса осповакцины и других поксвирусов, а также некоторых иридовирусов — *комплементарные цепи ДНК ковалентно соединены между собой*, причем как на самом конце молекулы, так и на непосредственно прилегающих к нему участках имеется несколько неспаренных (некомплементарных) нуклеотидов; кроме того, имеется инвертированный концевой повтор значительной длины. Происхождение и значение всех этих особенностей концевых структур станут понятными после рассмотрения способов репликации геномов соответствующих вирусов.

Известна группа вирусов с *линейным одноцепочечным* ДНК-геномом — это парвовирусы. Молекулы ДНК парвовирусов содержат 4,5—5,4 т. н. и характеризуются *самокомплементарностью концевых последовательностей*. Другими словами, концы такой молекулы способны образовывать элементы вторичной структуры («шпильки») с прилегающими участками той же молекулы (см. с. 268).

Различия между вирусными геномами касаются не только формы молекул ДНК. У некоторых вирусов — пока это известно только для вирусов прокариот, т. е. для фагов, — в состав ДНК входят *необычные и модифицированные основания*. Классический пример: ДНК Т-четных фагов вместо цитозина содержит 5-оксиметилцитозин (ОМЦ), причем к оксигруппе этого основания могут быть присоединены один или два остатка глюкозы (с образованием соответственно гликозил-ОМЦ и гентибиозил-ОМЦ). Геномы некоторых других фагов также содержат необычные основания. В большинстве случаев модифицированным оказывается пиримидин (тимин или цитозин). Ни одна из указанных модификаций не нарушает способности оснований вступать в стандартные уотсон-криковские взаимодействия.

Инициация цепи ДНК

Как уже отмечалось (см. гл. II) потребность в затравке — универсальное свойство ДНК-полимераз. Характер затравки и способ ее образования в значительной степени определяют своеобразие индивидуальных вирусных систем.

а. Инициация на одонитевых матрицах

Следует различать инициацию *на внутренних участках* одонитевой матрицы и *на ее концах*.

При попадании в клетку кольцевого одонитевого ДНК-генома инициация синтеза комплементарной цепи неизбежно должна проходить на внутренних участках матрицы — концов у молекулы

просто нет. Поучительно посмотреть, как это происходит при синтезе (—) цепи на (+) матрице ДНК фагов М13, G4 и фХ174. Во всех трех случаях образуется затравка, комплементарная единственному участку родительского генома, но эта цель достигается тремя различными путями.

У фага М13 олигорибонуклеотид-затравка образуется при помощи ДНК-зависимой РНК-полимеразы клетки-хозяина (*E. coli*). У фага G4 синтез затравки катализирует фермент праймаза — продукт клеточного гена *dna G*. Наконец, в случае фага фХ174 для образования затравки при синтезе (—) цепи ДНК используется многокомпонентная система, в которую наряду с праймазой — продуктом гена *dna G* — входит по крайней мере еще шесть клеточных белков: полипептиды *n*, *n'*, *n''* и продукты генов *dna B*, *dna C* и *dna T* (полипептид *i*). Эти белки формируют специальную структуру — праймосому, которая и осуществляет синтез затравки (см. гл. II, раздел 5). Собственно полимеразной активностью в праймосоме обладает праймаза.

Во всех трех случаях первая затравка для синтеза (—) цепи образуется на совершенно определенном участке фагового генома, разном у разных фагов. Соответствующий участок, очевидно, содержит сигналы в виде последовательности нуклеотидов и элементов вторичной структуры, которые специфически узнаются соответственно РНК-полимеразой (в ДНК фага М13), праймазой (у G4) и полипептидом *n'* (у фХ174). Подчеркнем, что в разобранных случаях матрицей для образования затравки является не «голая» ДНК, а дезоксирибонуклеопротеид, образованный в результате взаимодействия ДНК с ДНК-связывающим белком *E. coli*.

При синтезе полной (—)цепи ДНК фагов М13 и G4 используется единственная затравка, образованная так, как только что было описано. Однако при образовании (—)цепи ДНК фага фХ174 возникают множественные затравки. Дело в том, что праймосома обладает способностью передвигаться по матрице в энергозависимой реакции. Синтезировав затравку в первом (уникальном) месте, праймосома передвигается по матрице в 5'→3'-направлении, т. е. в сторону, противоположную удлинению затравки. Встретив «подходящую» нуклеотидную последовательность, праймосома катализирует образование новой затравки.

Такие же или сходные механизмы используются и для образования затравок на одноцепочечных участках, временно возникающих в районе репликационной вилки при полуконсервативной репликации двуцепочечных ДНК.

Отметим, что некоторые вирусные геномы кодируют собственные вирус-специфические праймазы (например, у фагов Т4 и Т7).

Что касается инициации на концах одноцепочечной ДНК-матрицы, то тут известны два механизма. Первый — *самозатравочный* — реализуется, например, у парвовирусов, 3'-конец генома которых, как уже известно, способен взаимодействовать с соседней

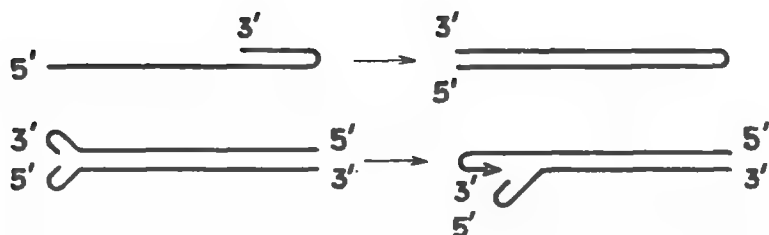


Рис. 135. Самозатравочный механизм синтеза цепи ДНК на одноцепочечной (сверху) и двухцепочечной (снизу) матрицах

[В последнем случае необходима предварительная перестройка конформации матричной молекулы — образование концевых внутрицепочечных «шпильек» («заячьих ушей»)]

комплементарной последовательностью, образуя «шпильку» (рис. 135). Такая структура имеет все качества, необходимые для функционирования в качестве затравки, и поэтому потребность в синтезе какой-либо иной затравки не возникает. Следовательно, в этом случае в качестве затравки используется не РНК, а ДНК.

При другом способе терминальной инициации роль затравки выполняет белок, точнее — *ковалентное соединение белка с нуклеотидом*. Такое соединение возникает в результате образования фосфодиэфирной связи между 5'-гидроксилом дезоксирибонуклеотида (например, dCMP) и гидроксилом оксиаминокислоты (например, серина) специального, так называемого терминального белка. В изученных вирусных системах терминальный белок — это всегда вирус-специфический (т. е. закодированный в вирусном геноме) полипептид, и фермент, осуществляющий присоединение нуклеотида, также всегда имеет вирус-специфическую природу. Нуклеотид-белковый комплекс взаимодействует с 3'-концом одноцепочечной вирусной ДНК-матрицы; при этом нуклеотид, входящий в комплекс с терминальным белком, комплементарен 3'-концевому нуклеотиду матрицы и служит затравкой, к которой присоединяются последующие нуклеотиды (рис. 136). Ясно, что к 5'-концу синтезированной таким образом цепи ДНК будет ковалентно присоединен белок. Рассмотренный способ инициации цепи ДНК реализуется, например, у аденовирусов и у фага $\phi 29$, у которых одонитевые ДНК-матрицы образуются в процессе репликации двунитевого генома (см. с. 267).

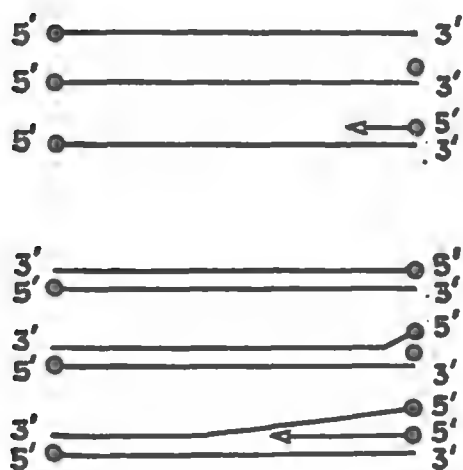


Рис. 136. Инициация синтеза цепи ДНК при помощи терминального белка-затравки на одноцепочечной (сверху) и двухцепочечной (снизу) матрицах

Кружок — терминальный (концевой) белок или нуклеотид-белковый комплекс

Кружок — терминальный (концевой) белок или нуклеотид-белковый комплекс

В основе некоторых способов инициации цепи ДНК на дуплексной матрице лежат те же механизмы, о которых шла речь в предыдущем разделе. Прежде всего это относится к *терминальной инициации*, для осуществления которой могут использоваться как *самозатра- вочный механизм*, так и *нуклеотид-белковая затравка*. В первом случае предварительно необходимо перестроить двухспиральный «тупой» конец в структуру типа «заячьи уши» (см. рис. 135), как это, по-видимому, происходит у репликативной формы парвови- русной ДНК. Поскольку спонтанная перестройка такого рода в изолированных молекулах ДНК крайне маловероятна по энерге- тическим соображениям, постулируется, что в ее осуществлении принимают участие какие-то (пока не идентифицированные) белки. Рис. 136 дает представление о том, как инициацию цепи ДНК на двухнитевой матрице можно обеспечить при помощи нуклеотид- белковой затравки. Вполне возможно, что и в этом случае некие белки способствуют расплетанию (или по крайней мере дестабили- зации) концевых участков дуплексного генома.

Что касается инициации *на внутренних участках* двухнитевой матрицы, то здесь также нужно различать два основных способа. Во-первых, *первичная РНК-затравка может быть образована прай- мазой* (или — реже — ДНК-зависимой РНК-полимеразой). Од- нако синтез затравки возможен только в том случае, если матрица соответствующим образом подготовлена. Подготовка включает взаимодействие между вирус-специфическими белками, регулирую- щими инициацию раунда репликации, и специфическими участками инициации репликации *ori* (от англ. origin — начало) в молекуле ДНК. Например, с участком *ori* в ДНК фага λ первично взаимо- действует фагоспецифический белок — О, с белком О взаимодей- ствует другой фагоспецифический полипептид — белок Р, который в свою очередь образует комплекс с одной из клеточных хеликаз — продуктом гена *dna B*.

Для образования первой затравки на молекуле ДНК SV40 необходимо присоединение к *ori* вирус-специфического белка — так называемого большого Т-антигена, который выполняет функ- ции хеликазы. Взаимодействие между *ori* и специфическими бел- ками создает необходимые условия для синтеза затравки фермен- тами, которые умеют это делать, обычно праймазой. Однако в некоторых системах (в частности, у того же фага λ) требуется дополнительное активирование *ori*. Эта цель может достигаться, например, тогда, когда в участке *ori* происходит транскрипция. Для такой *транскрипционной активации* важен именно сам акт транскрипции, а не ее продукты — РНК или белки. Считается, что в процессе транскрипции ослабляется связь между комплемен- тарными цепями; когда такое ослабление захватывает участок *ori*, он становится более доступным для праймазы.

Вторая группа способов внутренней инициации связана с использованием *разрывов, или брешей*, в одной из цепей двухнитевой ДНК. Эти «перерывы непрерывности» могут возникать в результате действия тех или иных нуклеаз на ДНК-матрицу; появление разрывов обычно сопровождает процесс рекомбинации. Так или иначе свободный 3'-ОН-конец цепи ДНК может быть использован в качестве затравки для дальнейшего удлинения этой цепи.

Элонгация цепи ДНК

Элонгация (удлинение) цепи ДНК осуществляется ДНК-зависимыми ДНК-полимеразами. В этой реакции участвуют также и вспомогательные белки, наборы которых могут различаться в разных системах и на разных этапах репликации одного и того же генома. В частности, различны эти наборы при синтезе ДНК на однонитевой матрице (или, как говорят, при «репарационном» синтезе) и на двухнитевой матрице (при синтезе с вытеснением цепи). В первом случае важным вспомогательным участником реакции являются ДНК-связывающие белки, которые превращают матрицу в дезоксирибонуклеопротейд. При этом исчезают многие из элементов вторичной структуры матрицы, она как бы «выпрямляется», что облегчает поступательное и процессивное движение ДНК-полимеразы. Сходную роль — помощь ДНК-полимеразе в преодолении «препятствий», в частности «шпильчатых» структур на матрице, — могут играть и другие дополнительные (в том числе и вирус-специфические) репликационные белки.

Еще разнообразнее наборы белков, участвующие в синтезе ДНК на двухнитевых матрицах. В этом случае помимо уже перечисленных, требуются, в частности, хеликазы, способствующие расплетанию родительского дуплекса в области репликационной вилки (см. гл. II), набор ферментов, необходимых для синтеза «отстающей» цепи (праймазы; ферменты, удаляющие РНК-затравку; ДНК-лигазы, сшивающие фрагменты Окадаки), а также — часто — топоизомеразы, снимающие избыточное внутримолекулярное напряжение, возникающее в результате расплетания матричного дуплекса. В общем, процесс элонгации при репликации вирусных ДНК-геномов не отличается принципиально от этого процесса при синтезе клеточных ДНК. Единственно, что следует отметить, — это использование (в некоторых системах) вирус-специфических репликационных белков, которые по своей функции аналогичны белкам, имеющимся в незараженной клетке.

Основные схемы репликации

Для понимания природы разнообразия в способах репликации вирусных ДНК-геномов важно ясно представлять себе трудность полного воспроизведения линейной молекулы ДНК. Эта труд-

ность — прямое следствие двух универсальных (не знающих исключений) свойств ДНК-полимераз: 1) полярности синтеза (всегда от 5'- к 3'-концу вновь строящейся цепи или, что то же самое, от 3'- к 5'-концу матричной цепи) и 2) потребности в затравке. Например, в случае использования РНК-затравок невозможно (если не прибегать к специальным ухищрениям) получить ДНК-копию 3'-конца матричной цепи (рис. 137). Предположим, что на этом месте синтезирована РНК-затравка. Трудность в том, чтобы заполнить брешь, которая возникает после удаления затравки — ведь ДНК-полимераза «заднего хода» не имеет. Оказывается, что существует несколько — всего несколько — основных способов, обеспечивающих полное воспроизведение ДНК-геномов. Проиллюстрируем эти способы кратким описанием конкретных вирусных систем.

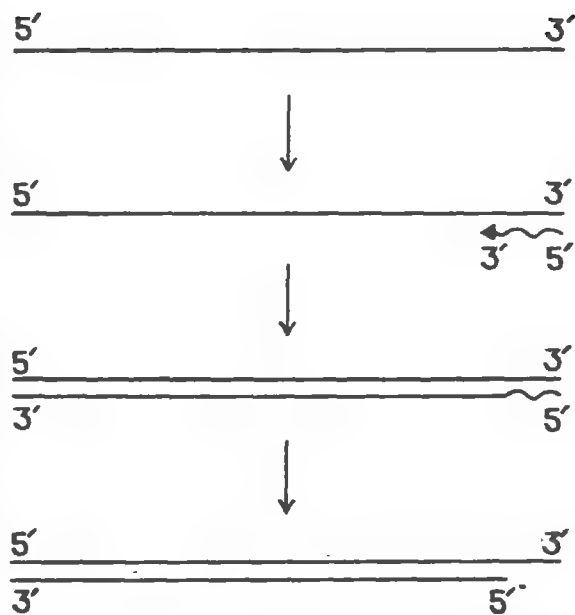


Рис. 137. Проблема, связанная с полным воспроизведением линейной молекулы ДНК

Прямые линии — ДНК; волнистые — РНК. После удаления РНК-затравки остается брешь, которую трудно заполнить

а. Репликация генома аденовирусов (терминальная инициация с использованием нуклеотид-белковой затравки)

Общая схема репликации двухнитевой аденовирусной ДНК представлена на рис. 138. К синтезированному в зараженной клетке вирус-специфическому терминальному белку (точнее — к его более крупному предшественнику) ковалентно присоединяется нуклеотид, и этот нуклеотид-белковый комплекс выполняет роль затравки. Синтез новой цепи может начаться на любом конце генома, так как в силу наличия инвертированных концевых повторов оба конца молекулы аденовирусной ДНК идентичны. Синтез дочерней цепи сопровождается вытеснением одной из родительских цепей. После полного вытеснения этой цепи высвобождается ее 3'-конец, к которому присоединяется новая нуклеотид-белковая затравка. Синтез комплементарной цепи происходит теперь на однострочной матрице, т. е. по «репарационному» типу. Результат этих двух актов — синтеза с вытеснением на родительском дуплексе и «репарационного» синтеза на вытесненной цепи — полностью эквива-

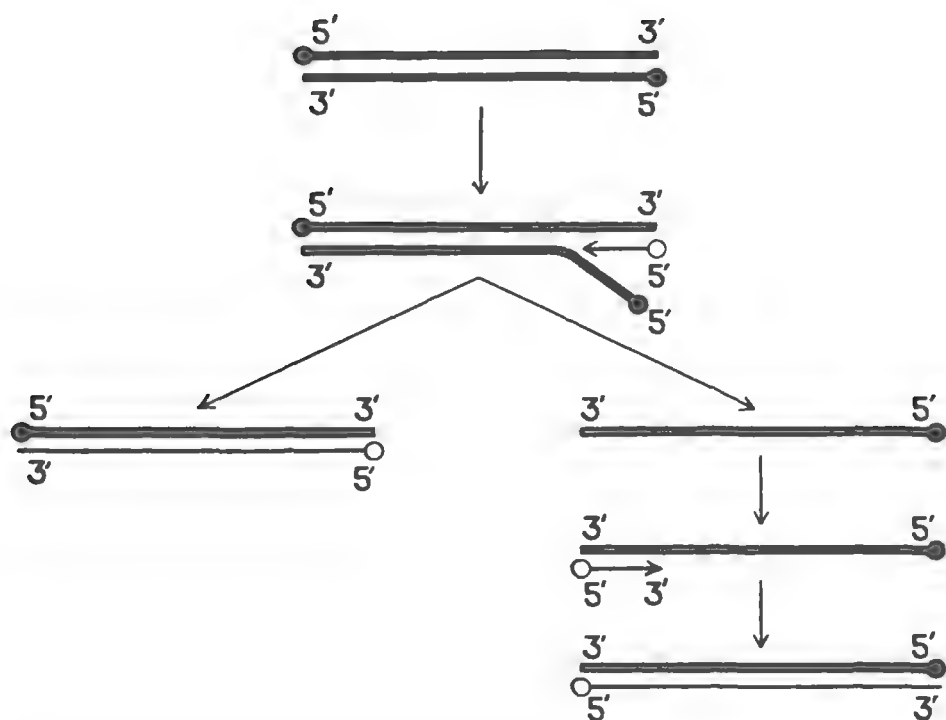


Рис. 138. Схема репликации генома аденовируса.

Жирные линии — родительские цепи ДНК, тонкие — дочерние; кружки — молекулы терминального белка (черные — родительские; светлые — дочерние)

лентен полуконсервативной репликации двухнитевой молекулы ДНК: возникают две дочерние двухнитевые молекулы, каждая из которых физически наследует по одной из цепей родительского генома.

По такой же общей схеме реплицируется геном фага ф29.

б. Репликация генома парвовирусов (терминальная инициация при помощи самозатравочного механизма)

Парвовирусы — мелкие ДНК-содержащие вирусы животных — имеют в качестве генома однонитевую молекулу ДНК, оба конца которой способны формировать «шпилечные» структуры благодаря наличию самокомплементарных последовательностей. В качестве примера рассмотрим схему репликации ДНК аденоассоциированного вируса (репродукция этого вируса требует, чтобы в той же клетке размножался вирус-помощник, каковым может быть, в частности, аденовирус) (рис. 139).

После попадания вирусного генома в клетку его 3'-концевой сегмент образует как бы Т-образную двухнитевую структуру (из-за соответствующего расположения самокомплементарных последовательностей); эта структура и выполняет роль затравки. В ре-

зультате «репарационного» синтеза комплементарной цепи возникает дуплекс, обе цепи которого на одном из концов ковалентно соединены между собой. Нетрудно видеть, что 3'-концевой сегмент родительского генома (длина его 125 нуклеотидов) при этом в качестве матрицы не использовался и, следовательно, полного воспроизведения генома еще не произошло. Затем вирус-специфический фермент вносит разрыв в родительскую цепь как раз на границе между реплицированным и нереплицированным участками (т. е. между 125-м и 126-м нуклеотидами, если считать от 3'-конца исходного генома). Концевые 125 нуклеотидов родительского генома становятся при этом как бы частью комплементарной цепи. Возникший таким образом 3'-конец комплементарной цепи используется в качестве матрицы для «регенерации» недостающей концевой последовательности родительской цепи. Как понятно из схемы, «регенерация» сопровождается инверсией: если исходную концевую структуру родительской цепи обозначить $1 \rightarrow 125$, то после завершения процесса она приобретает вид $125' \rightarrow 1'$, где штрих у цифры обозначает комплементарность нуклеотида (т. е. нуклеотид $1'$ комплементарен нуклеотиду 1). В результате этих реакций возникает двухнитевая — *репликативная* — форма вирусной ДНК.

Далее следует перестройка одного из концов этой молекулы с образованием «заячьих ушей». При этом возникает затравка, с которой начинается синтез цепи, имеющей полярность, например, родительского генома (полярность будет, естественно, зависеть от того, на каком из концов репликативной формы образуются «заячьи уши»). В таком случае синтез новой цепи будет сопровождаться вытеснением исходной родительской цепи из дуплекса; другим продуктом реакции будет двухнитевая репликативная форма, у которой комплементарные цепи опять ковалентно связаны между собой, но на этот раз уже на другом конце молекулы. Затем происходит «регенерация» и инверсия 3'-конца комплементарной цепи. Образующаяся при этом молекула репликативной формы будет использована в качестве матрицы для дальнейшего синтеза вирусной ДНК, а вытесненная из дуплекса одонитевая молекула имеет выбор: она может либо также вступить в репликативный пул, либо войти в состав дочерней вирусной частицы (второй исход возможен только после накопления в зараженной клетке достаточного количества структурных вирусных белков).

Благодаря инверсиям концевые последовательности дочерних геномов могут различаться между собой. Поскольку каждый конец молекулы может находиться в одной из двух ориентаций, причем ориентация одного конца не зависит от ориентации другого, и учитывая, что в состав популяции аденоассоциированных вирусов входят примерно поровну частицы с геномами положительной и отрицательной полярности, зараженная клетка генерирует восемь типов дочерних геномов; все эти геномы в биологическом отношении эквивалентны.

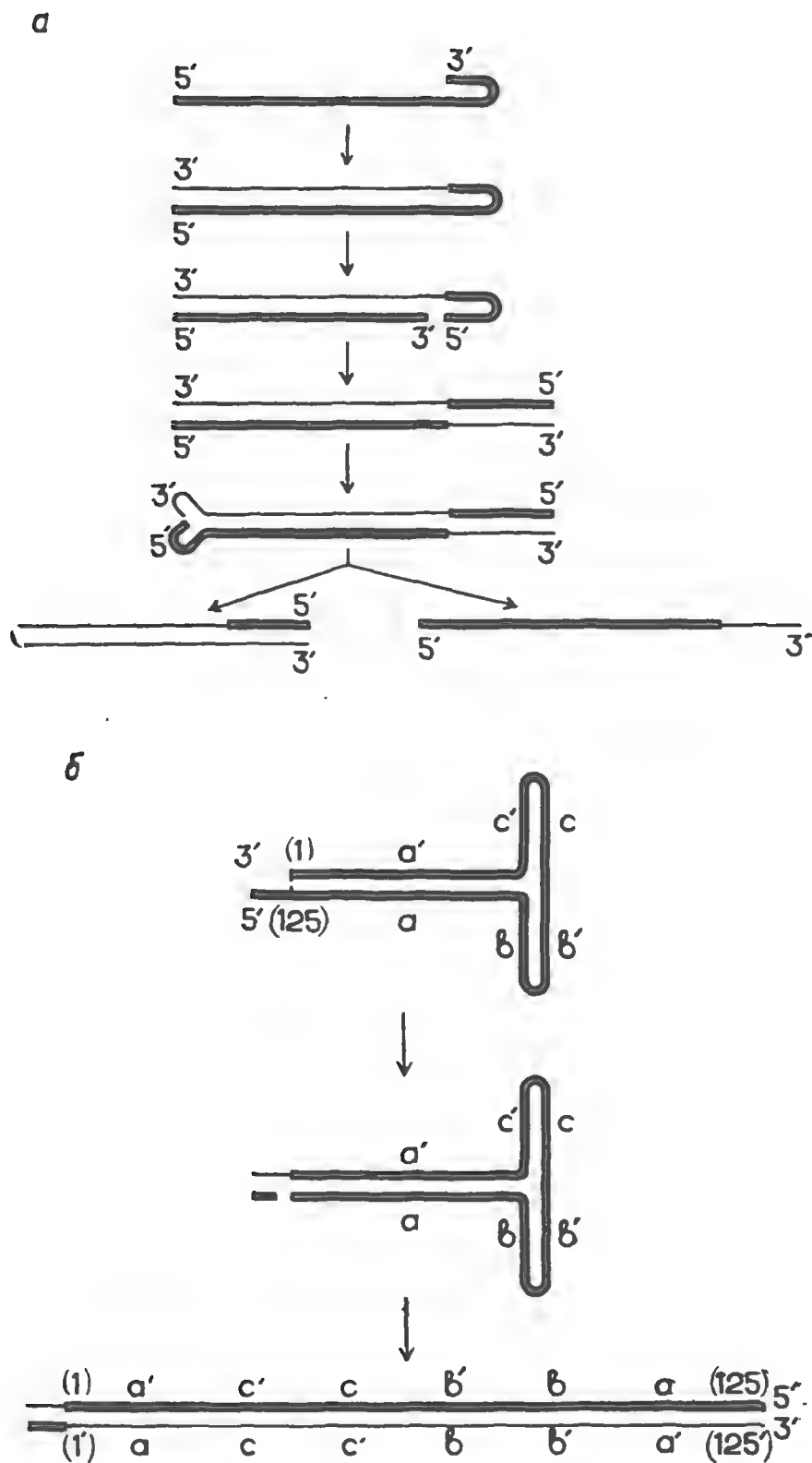


Рис. 139. Схема репликации генома аденоассоциированного вируса:

a — общая последовательность стадий: самозатравочный синтез комплементарной цепи, эндонуклеолитический разрыв, «регенерация» 3'-конца родительской цепи, образование «заячьих ушей», синтез второй дочерней цепи, сопровождающийся вытеснением родительской цепи из дуплекса; *б* — детальное описание событий, происходящих на 3'-конце родительской цепи и приводящих к появлению инверсии; жирные линии — родительские последовательности, тонкие — дочерние; буквы — сегменты ДНК; цифры в скобках — номера нуклеотидов

в. Репликация генома паповавирусов (репликация по схеме Кэрнса)

Наиболее изученным представителем паповавирусов является онкогенный обезьяний вирус 40 (SV40). После взаимодействия вирус-специфической хеликазы — большого Т-антигена — с участком *ori* в кольцевой двухнитевой ковалентно-непрерывной молекуле ДНК SV40 праймаза образует затравку сначала на одной из родительских цепей, а потом и на другой. Соответственно возникает две лидирующие (и две отстающие) цепи и две репликативные вилки, удаляющиеся друг от друга в результате двунаправленной репликации (рис. 140). Ферментативный аппарат, необходимый для элонгации, поставляется клеткой, если не считать хеликазной активности Т-антигена. Если выделить реплицирующиеся молекулы — их иногда называют молекулами θ -типа или θ -молекулами, — то можно убедиться, что обе родительские цепи в них сохраняют ковалентную непрерывность. Однако это обстоятельство должно создавать серьезную проблему по мере расхождения репликативных вилок друг от друга. Действительно, синтез дочерних цепей сопровождается расплетанием родительского дуплекса. Расплетание кольцевого дуплекса в одном месте будет вести к резкому повышению внутримолекулярных напряжений в нереплицированных участках молекулы. Напряжение могло бы вести к сверхспирализации ДНК (см. гл. I, раздел 2). Расплетание всего 10 п. н. дуплекса, как известно, приводит к появлению одного сверхспирального витка. Соответственно репликация половины молекулы должна была бы сопровождаться появлением около 250 сверхспиральных витков (напомним что всего в геноме SV40 около 5 т. п. н.). Этого не происходит, да и не может произойти — просто нет места. Сбрасывание внутримолекулярных напряжений обеспечивается действием клеточных топоизомераз, в первую очередь ДНК-топоизомеразы I, которая вносит временные одноцепочечные разрывы в родительские

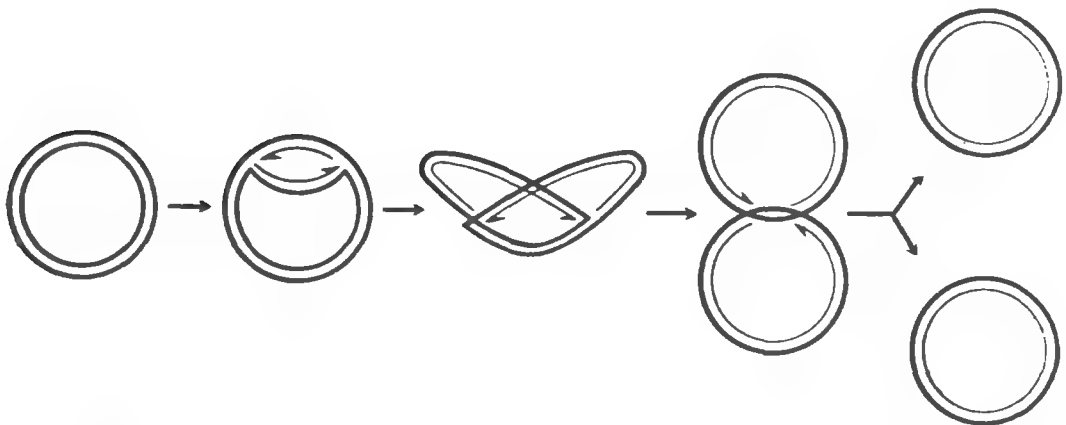


Рис. 140. Схема репликации генома SV40 (схема Кэрнса)

Жирные линии — родительские цепи ДНК; тонкие — дочерние

нити. Схема репликации ДНК, которая здесь излагается, была впервые предложена Кэрнсом для объяснения способа репликации кольцевой хромосомы кишечной палочки.

На завершающих стадиях раунда репликации образуются две почти полные кольцевые молекулы, содержащие по одной недостроенной дочерней цепи и по одной полной (кольцевой) родительской цепи. При этом родительские цепи остаются зацепленными друг за друга (рис. 140). Затем происходит достаривание дочерних цепей и лигирование их концов. Образовавшиеся два ковалентно-непрерывных кольцевых дуплекса расцепляются при помощи клеточной ДНК-топоизомеразы II. Расцепление колец может происходить и до завершения синтеза дочерних молекул.

г. Репликация генома
с односторонней кольцевой ДНК
(репликация по схеме
разматывающегося рулона)

Как уже было сказано, ряд фагов (φX174, G4, M13 и др.) имеют одностороннюю кольцевую геном. Вскоре после попадания такого генома в клетку он превращается в кольцевой ковалентно-непрерывный дуплекс (или, как говорят, в *репликативную форму*). Это превращение включает ряд стадий: 1) образование затравки; 2) элонгацию комплементарной цепи, осуществляемую клеточной ДНК-полимеразой III; 3) удаление РНК-затравки, которое производится, по-видимому, за счет 5'-экзонуклеазной активности клеточной ДНК-полимеразы I; 4) достраивание комплементарной цепи; 5) лигирование концов комплементарной цепи ДНК-лигазой и 6) внесение сверхспиральных витков в ковалентно-непрерывный дуплекс при помощи гиразы. Обратим внимание, что все ферменты, обеспечивающие перевод родительского генома в репликативную форму, имеют клеточное происхождение.

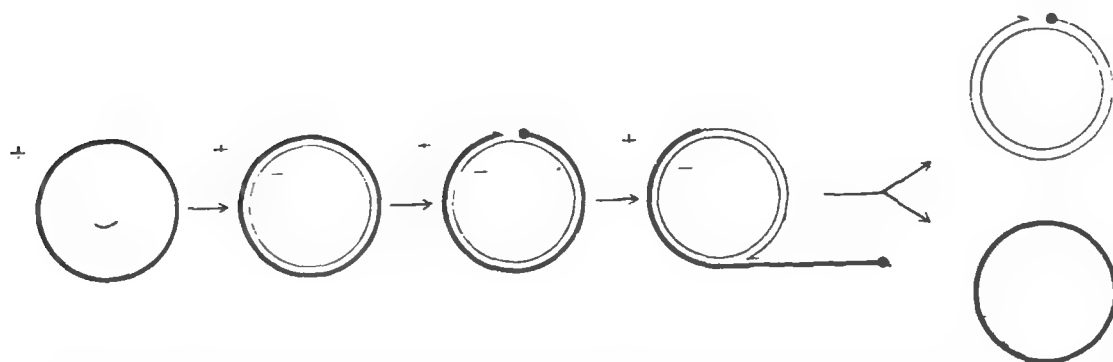


Рис. 141. Схема репликации генома фага φX174 (схема разматывающегося рулона)

Жирные линии — родительская цепь ДНК; тонкие — дочерние цепи; стрелка — растущий 3'-конец (÷) цепи, кружок — белок А, ковалентно присоединенный к 5'-концевому нуклеотиду

Дальнейшие события разберем на примере фага ϕ X174 (рис. 141); у некоторых других фагов этой группы репликация генома имеет особенности, которые, впрочем, не изменяют общую схему. Фагоспецифический белок — продукт гена А — вносит одноцепочечный разрыв в уникальное место родительской цепи в молекуле репликативной формы; одновременно он соединяется ковалентно (при помощи фосфодиэфирной связи) с генерируемым при разрыве 5'-концевым нуклеотидом этой цепи. Возникающий в месте разрыва нефосфорилированный 3'-концевой нуклеотид представляет собой *затравку*. Впрочем, эта затравка может быть использована клеточной ДНК-полимеразой III для элонгации 3'-конца родительской (т. е. (+)) цепи только при условии, что предсуществующий 5'-конец этой цепи будет вытеснен из дуплекса и тем самым обнажится однонитевая (—) матрица. Такое вытеснение осуществляется одной из клеточных хеликаз — продуктом гена *rep*.

Удлинение 3'-конца (+)цепи, сопровождающееся вытеснением ее 5'-конца, приводит к образованию промежуточных реплицирующихся молекул, напоминающих по форме греческую букву σ , — молекул σ -типа, или просто σ -молекул. Соответствующий способ синтеза ДНК на кольцевой матрице называют способом *разматывающегося рулона*. Следует заметить, что в реплицирующихся молекулах ДНК фага ϕ X174 белок А, находящийся на 5'-конце вытесняемой (+)цепи, по-видимому, удерживается в области репликационной вилки, возможно, за счет взаимодействия с хеликазой Rep. Поэтому форма реплицирующейся молекулы в этом случае скорее напоминает цифру 8.

После того как репликационная вилка совершит чуть больше полного оборота по (—)матрице и будет синтезировано чуть больше полного эквивалента вирусного генома, вытесненная (+)цепь ДНК замыкается в кольцо. Схематически этот интересный процесс можно представить следующим образом (рис. 142). Белок А, находящийся в области репликационной вилки на 5'-конце вытесненной цепи, узнает «свое» место на границе между 3'-концом полностью вытесненного генома [(+)цепи] и 5'-концом следующего генома, еще находящегося в дуплексе. «Узнав» это место, белок А вносит в него разрыв; при этом он отсоединяется от уже вытесненной цепи и присоединяется к 5'-концевому нуклеотиду (+)цепи, находящейся в дуплексе. Одновременно 5'- и 3'-концы вытесненной (+)цепи лигируются, и тем самым она превращается в кольцо. Все эти реакции осуществляются одномоментно (точнее — сопряженно) и без затраты внешней энергии.

Таким образом, вновь синтезированная (+)цепь становится компонентом репликативной формы, а предсуществовавшая (родительская) кольцевая (÷)цепь оказывается в свободном виде. По уже разобранному механизму она превращается в молекулу репликативной формы, которая вовлекается в синтез новых (+)колец; эти кольца опять превращаются в дуплексы и т. д. Все это при-

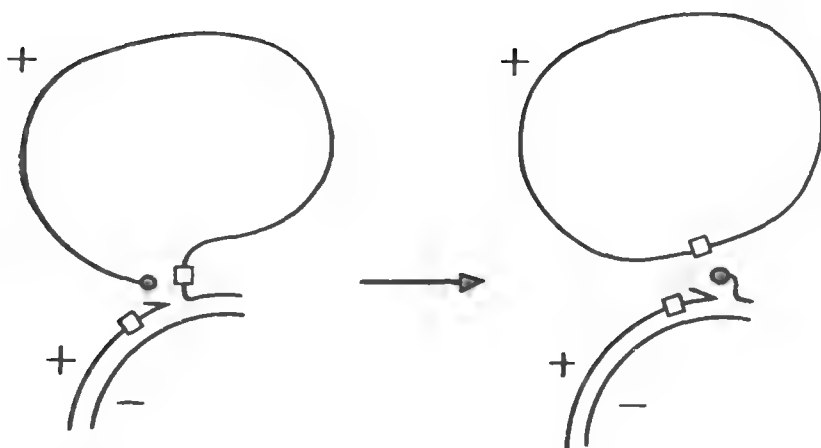


Рис. 142. Схема терминации синтеза (+) цепи фага ΦX174

Слева — фрагмент реплицирующейся кольцевой молекулы, в которой прошел полный раунд синтеза (+)цепи; стрелка — растущий 3'-конец (+)цепи; кружок — белок А, ковалентно присоединенный к 5'-концевому нуклеотиду; квадрат — последовательность, узнаваемая белком А [ori (+)цепи]; справа — продукт терминации: кольцевая (+)цепь и молекула репликативной формы, начинающая новый раунд репликации

водит к накоплению в зараженной клетке нескольких десятков молекул репликативной формы ДНК.

На молекулах репликативной формы ДНК происходит синтез не только (+)цепей ДНК, но и вирус-специфических мРНК. Следует сказать, что синтез мРНК должен предшествовать появлению новых молекул (+)цепей ДНК, так как без вирус-специфических мРНК в зараженной клетке не может появиться белок А. Трансляция фаговых мРНК приводит к накоплению вирус-специфических белков, в том числе и структурных, которые — при достаточной концентрации — начинают превращаться в сложные структуры — предшественники вирусного капсида. Генерируемые на этой стадии (+)кольца в результате специфических взаимодействий с белками фага вовлекаются в процесс сборки вириона. Тем самым предотвращается ставший уже ненужным переход (+)цепей в репликативную форму.

д. Репликация генома фага лямбда (репликация по схеме вторичного разматывающегося рулона)

После попадания в клетку линейной двухнитевой ДНК фага λ она циркуляризуется за счет взаимодействия между липкими концами и превращается в ковалентно-непрерывную сверхспирализованную форму. Раунд репликации начинается с образования затравки в специфическом *ori*-участке; распознавание этого участка осуществляется фагоспецифическими белками, а собственно синтез затравки производит клеточная праймаза, причем требуется

транскрипционная активация (см. с. 265). Вначале репликация идет по схеме Кэрнса: обнаруживаются θ -молекулы, в которых две репликационные вилки движутся в противоположных направлениях. Как всегда, движение репликационной вилки, имеющей лидирующую и отстающую цепи, требует сложного ферментативного аппарата. Большинство компонентов этого аппарата (ДНК-полимераза, праймаза, лигаза, хеликазы и др.) при репликации генома фага λ поставляется клеткой, хотя определенную роль, по-видимому, играют и фагоспецифические белки О и Р.

Схема Кэрнса, однако, не является основным способом репликации ДНК фага λ . Очень быстро (может быть уже после первого раунда) θ -молекулы превращаются в σ -молекулы, т. е. приобретают форму, характерную для участников репликации по схеме разматывающегося рулона. Тем не менее между σ -молекулами репликационных систем, использующих классический механизм разматывающегося рулона (например, у фага $\phi\chi 174$), и σ -молекулами, образующимися на поздней стадии репликации ДНК фага λ , есть существенные различия. В первом случае 5'-конец «хвостовой» части молекулы имеет совершенно определенную структуру, так как он возникает в результате внесения разрыва в уникальное место кольцевого дуплекса. В случае же ДНК фага λ σ -молекулы могут иметь самые разнообразные концы. При классическом разматывающемся рулоне из дуплекса вытесняется всегда определенная цепь [(+)цепь у $\phi\chi 174$], что опять-таки связано с уникальностью разрыва, вносимого в дуплекс. В случае σ -молекул ДНК фага λ из дуплекса может вытесняться любая из двух комплементарных цепей. Наконец, ферментативное обеспечение репликации по схеме разматывающегося рулона имеет свои особенности (например, в случае фага $\phi\chi 174$ — потребность в хеликазе Per), которые не обнаруживаются при поздней репликации генома фага λ (где используется тот же набор ферментов, что и на ранней стадии).

Все эти различия становятся легко объяснимыми, если принять, что переход от ранней к поздней стадии репликации ДНК фага λ связан с внесением более или менее случайных разрывов в разные цепи θ -молекулы. Появление при этом молекулы σ -типа становится понятным из рис. 143; понятно также, почему при этом не появляются новые потребности в ферментативном обеспечении: на ранней и поздней стадиях функционируют, по существу, одни и те же репликационные вилки. Все это дает основание отличать механизм репликации ДНК фага λ от схемы классического разматывающегося рулона. Для удобства будем называть способ поздней репликации генома фага λ способом *вторичного разматывающегося рулона*.

Поздняя репликация ДНК фага λ ведет к образованию σ -молекул, «хвостовая» часть которых содержит множественные тандемные копии (*конкатемеры*) вирусного генома. На заключитель-

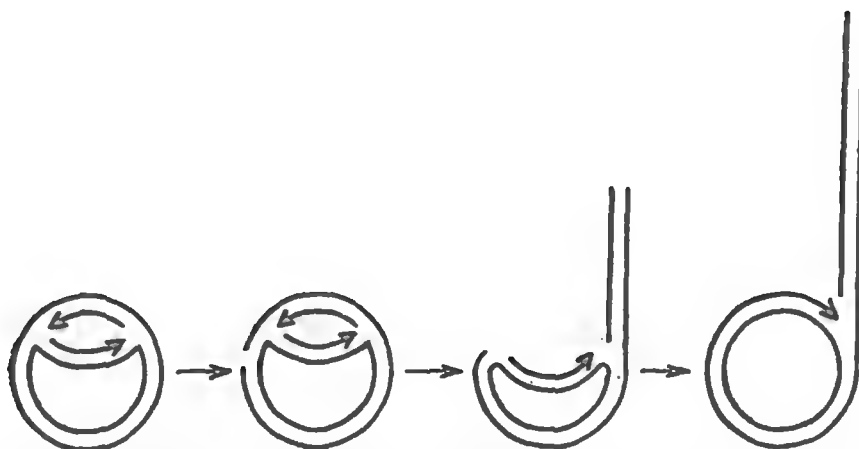


Рис. 143. Схема появления молекул σ -типа при репликации генома фага λ (схема вторичного разматывающего рулона)

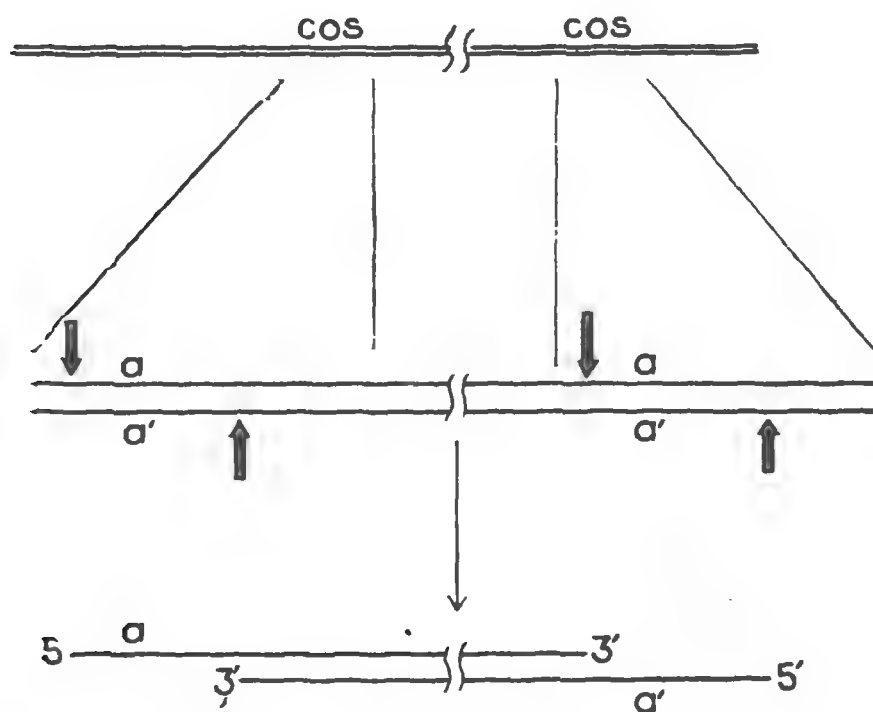


Рис. 144. Схема образования «липких» концов в молекуле ДНК фага λ

Сверху — участок конкатемерной фаговой ДНК; в середине (в большем масштабе) — два соседних участка *cos*; внизу — зрелая молекула; жирные стрелки — места одноцепочечных разрывов, вносимых матуразой; *a* и *a'* — последовательности, соответствующие «липким» концам

ной стадии эти конкатемеры превращаются в зрелые линейные мономеры с липкими концами. Процесс *созревания* ДНК сопряжен с ее упаковкой в предшественники фаговых головок. При этом активируется фагоспецифическая эндонуклеаза (в данном случае ее называют *терминазой* или *матуразой*), которая вносит ступенчатые разрывы в уникальные места — так называемые *cos*-участки — обеих цепей фаговой ДНК (рис. 144).

е. Репликация генома фагов Т-серии
(репликация с использованием
промежуточных конкатемерных форм)

Репликация ДНК ряда вирусов осуществляется при участии промежуточных конкатемерных форм. Конкатемеры могут возникать при синтезе ДНК на кольцевой матрице, например при репликации по схеме вторичного разматывающегося рулона. Однако они могут образовываться и без участия кольцевых молекул. Именно так обстоит дело у фагов Т-серии; при этом между системами репликации Т-нечетных фагов (например, Т7) и Т-четных (Т4) имеются заметные различия. Сначала рассмотрим более простую систему, которая реализуется у фага Т7 (рис. 145).

ДНК фага Т7 — линейная двухнитевая молекула (~ 40 т. п. н.) с прямым концевым повтором длиной 160 нуклеотидов. Инициация раунда репликации происходит внутри молекулы — на расстоянии, примерно соответствующем 15 % длины генома от одного из концов, условно называемого левым. Здесь имеются промоторы, которые узнаются фагоспецифической ДНК-зависимой РНК-полимеразой. Без транскрипции этого *ori*-района раунд репликации не начинается. Скорее всего РНК, образующаяся при транскрипции с этих промоторов, непосредственно используется в качестве за-

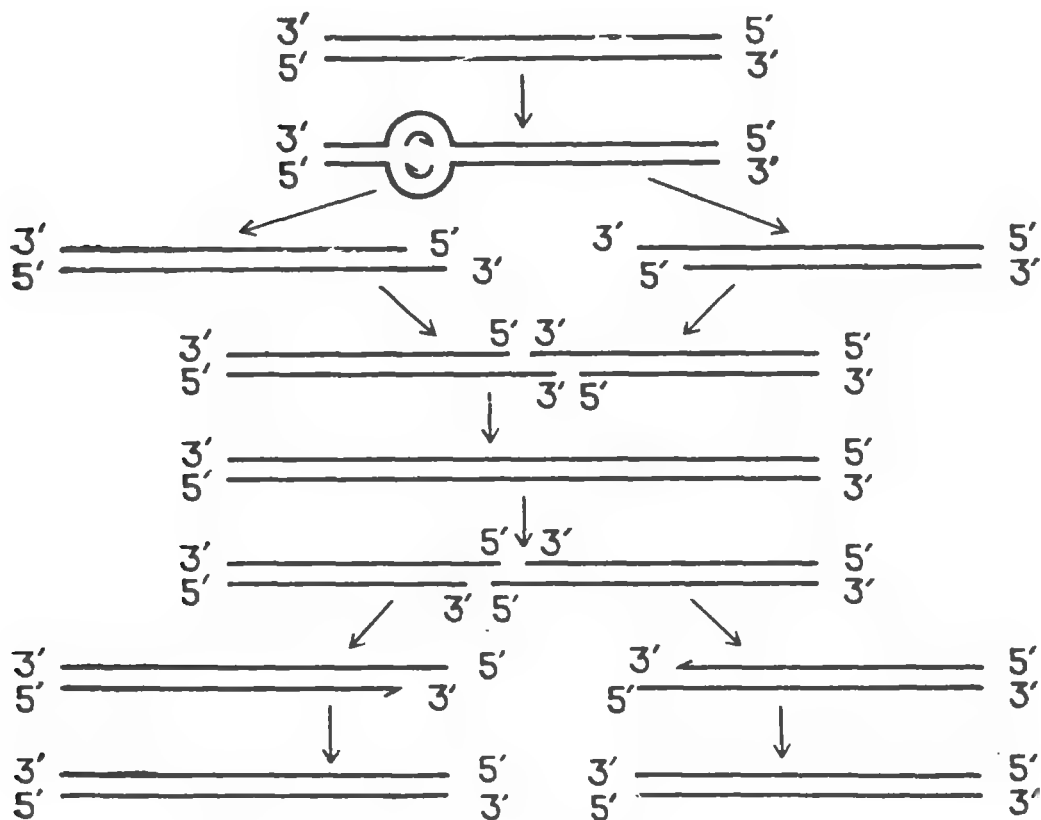


Рис. 145. Схема репликации генома фага Т7

травки для синтеза вирусной ДНК. Так или иначе внутренняя инициация происходит без разрыва родительских цепей ДНК. Возникшие две репликационные вилки, передвигаясь в противоположных направлениях, осуществляют полуконсервативную репликацию фагового генома (на ферментативном аппарате элонгации здесь останавливаться не будем; отметим лишь, что в данном случае он содержит преимущественно вирус-специфические белки). Первая стадия этого процесса заканчивается образованием двух дочерних дуплексов. Эти дуплексы имеют очень важную особенность — 5'-концы их вновь синтезированных цепей недостроены; другими словами, не произошло копирования 3'-концов родительских цепей. Такая ситуация неизбежно возникает при внутренней инициации на линейной ДНК-матрице. Причина нехватки понятна и является следствием двух универсальных свойств ДНК-полимераз: потребности в затравке и способности синтезировать цепочку ДНК только в одном направлении — от 5'-конца к 3'-концу.

Таким образом, один из 3'-концов вновь возникших дуплексов должен находиться в односторонней форме. Поскольку молекула ДНК фага Т7 имеет прямой концевой повтор, односторонние 3'-концы двух «сестринских» молекул должны быть взаимно комплементарны. Поэтому эти молекулы могут объединиться в димер, мономерные составляющие части которого на первых порах удерживаются за счет взаимодействия между комплементарными последовательностями, а затем скрепляются ковалентно. Сходным образом репликация димерной молекулы может привести к возникновению тетрамера и т. д.

На заключительной стадии из конкатемеров образуются полноценные мономерные геномы. Схематически это «созревание» можно представить следующим образом. Во внутренний участок конкатемерной молекулы — в районе, соответствующем концу генома, — вносится ступенчатый разрыв. При этом возникают мономерные дуплексы, у которых цепи имеют нехватку, но на этот раз на 3'-концах. Особых проблем с достраиванием таких концов не возникает. Можно отметить, что созревание геномов фага Т7 происходит сопряженно с их упаковкой в головку.

Способ репликации ДНК-генома фага Т4 (170 т. п. н.) отличается от вышеописанного в нескольких отношениях. Во-первых, у Т4 в этом процессе принимает участие большее число вирус-специфических белков, которые обеспечивают не только синтез аномальных (содержащих 5-оксиметилцитозин) нуклеотидов и пост-репликационное гликозилирование ДНК, но и выполняют множество парциальных реакций, необходимых непосредственно для синтеза молекулы ДНК. Во-вторых, инициация синтеза новых цепей ДНК на ранней и поздней стадиях репродукции фага Т4 происходит по-разному. На ранней стадии имеет место инициация на внутренних участках дуплекса (*ori*-сайтах). Определенную роль в этом процессе выполняет клеточная ДНК-зависимая РНК-

полимераза, которая обеспечивает либо образование затравки, либо транскрипционную активацию. Если не вдаваться в детали, ранняя инициация синтеза ДНК у Т4 схожа с таковой у фага Т7. Однако на поздних стадиях основной вклад в инициацию цепей ДНК Т4 вносит другой процесс, который называют *рекомбинационной инициацией*.

Суть его можно представить следующим образом (рис. 146). В результате первого раунда репликации возникают два дочерних дуплекса с одноцепочечными 3'-концами (как и при репликации генома фага Т7). Но геном Т4 характеризуется кольцевыми перестановками (см. рис. 134). Поэтому если клетка заражена несколькими фаговыми частицами, то концевая последовательность одной молекулы фаговой ДНК будет соответствовать внутреннему участку другой молекулы. При помощи специальных фагоспецифических белков одноцепочечный 3'-конец первой молекулы может внедриться во внутренний район другой молекулы; в результате появляется затравка, способная обеспечить дальнейшее удлинение цепи. В конечном счете возникает молекула, которую можно рассматривать как разветвленный конкатемер. Отметим, что рекомбинационная инициация имеет место и тогда, когда клетка заражается единственной фаговой частицей (рекомбинация в этом случае происходит между одинаковыми дочерними молекулами.)

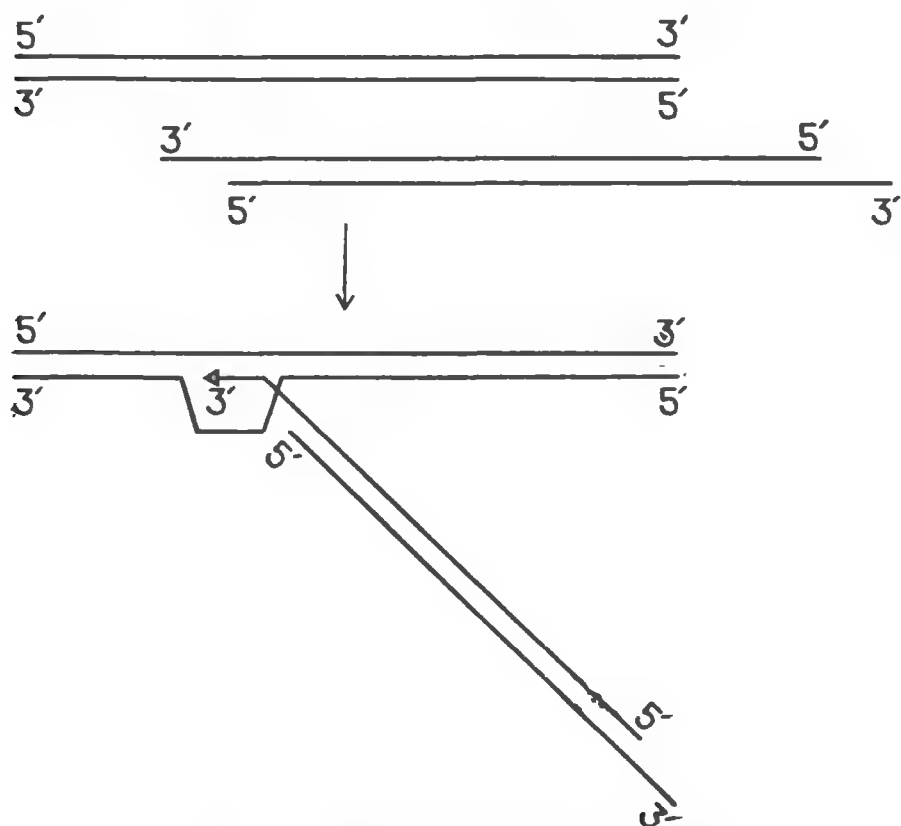


Рис. 146. Схема «поздней» (рекомбинационной) инициации цепей ДНК фага Т4

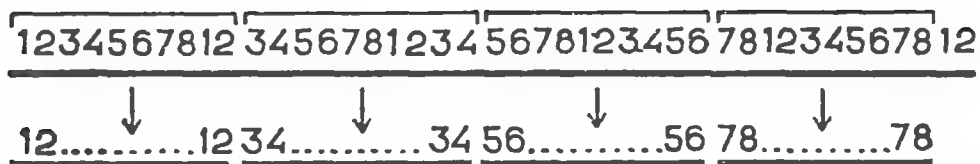


Рис. 147. Схема образования «зрелых» молекул ДНК фага Т4 (с концевыми перестановками и прямыми концевыми повторами) из конкатемера-предшественника

Цифры — сегменты фаговой ДНК; квадратные скобки — количество ДНК, вмещающееся в головку фага

Третье различие между системами репликации ДНК фагов Т4 и Т7 касается способа превращения конкатемера в зрелый мономерный геном. В первом случае длина сегмента ДНК, отрезаемого от конкатемера, задается не специфической нуклеотидной последовательностью (как у Т7), а вместимостью фаговой головки: конкатемерная молекула ДНК начинает упаковываться в головку, а когда головка заполнится, активируется эндонуклеаза, которая отщепляет оставшийся снаружи участок молекулы. Поскольку в головку помещается сегмент ДНК, превышающий по своим размерам уникальную последовательность вирусного генома, повторение актов упаковки и нарезания генерирует молекулы с кольцевыми перестановками и прямыми концевыми повторами (рис. 147). Отметим, что в фаговом геноме закодирован фермент, способствующий превращению разветвленных молекул ДНК в линейные.

Блочный принцип организации репликационных систем

Описанные выше схемы репликации ДНК-геномов включают почти весь набор основных элементов (блоков), из которых построены репликационные системы и других ДНК-содержащих вирусов (если не считать стоящие особняком способы репликации ДНК с участием механизмов обратной транскрипции; см. раздел 3 этой главы). Однако комбинироваться эти блоки могут в разных сочетаниях.

Удобно расчленить раунд репликации ДНК на три стадии: 1) переход родительского генома в репликативную форму; 2) собственно репликация репликативной формы и 3) переход репликативной формы в зрелый дочерний геном. Рассмотрим несколько вирусных систем, у которых синтез ДНК осуществляется при участии двухнитевых кольцевых молекул (рис. 148). Такие кольца — репликативные формы — могут возникать несколькими способами: путем синтеза комплементарной цепи на однонитевой кольцевой матрице (фаг ϕ X174), в результате спаривания липких концов (фаги λ , P2, P4), в результате сайт-специфической (фаг P1) или общей (фаг P22) внутримолекулярной рекомбинации между концевыми повторами и т. д. Наконец, в форме двухнитевого кольца

родительский геном может находиться исходно (SV40); в этом случае первая стадия репликационного цикла — переход родительского генома в репликативную форму — как бы отсутствует (впрочем, репликативной формой в этом случае часто называют только молекулы θ -типа).

Способ репликации репликативной формы не зависит от способа ее образования: так, кольцевые репликативные формы у фагов P2, P4 и λ возникают, как уже упоминалось, одним и тем же путем — в результате взаимодействия между липкими концами, однако в первом случае репликация осуществляется по схеме (первичного) разматывающегося рулона, во втором — по схеме Кэрнса, а в третьем — по схеме вторичного разматывающегося рулона. Точно так же способ созревания дочерних геномов не связан жестко со способом репликации. Например, дуплексы с липкими концами могут возникать как из конкатемеров (λ), так и из колец (P2 и P4); геномы, образующиеся в результате нарезания конкатемеров, могут иметь не только «липкие» (i -), но и «тупые» (P22) концы; они могут иметь или не иметь кольцевые перестановки и т. д.

Использование ограниченного числа стандартных блоков для осуществления каждой из трех последовательных стадий реплика-

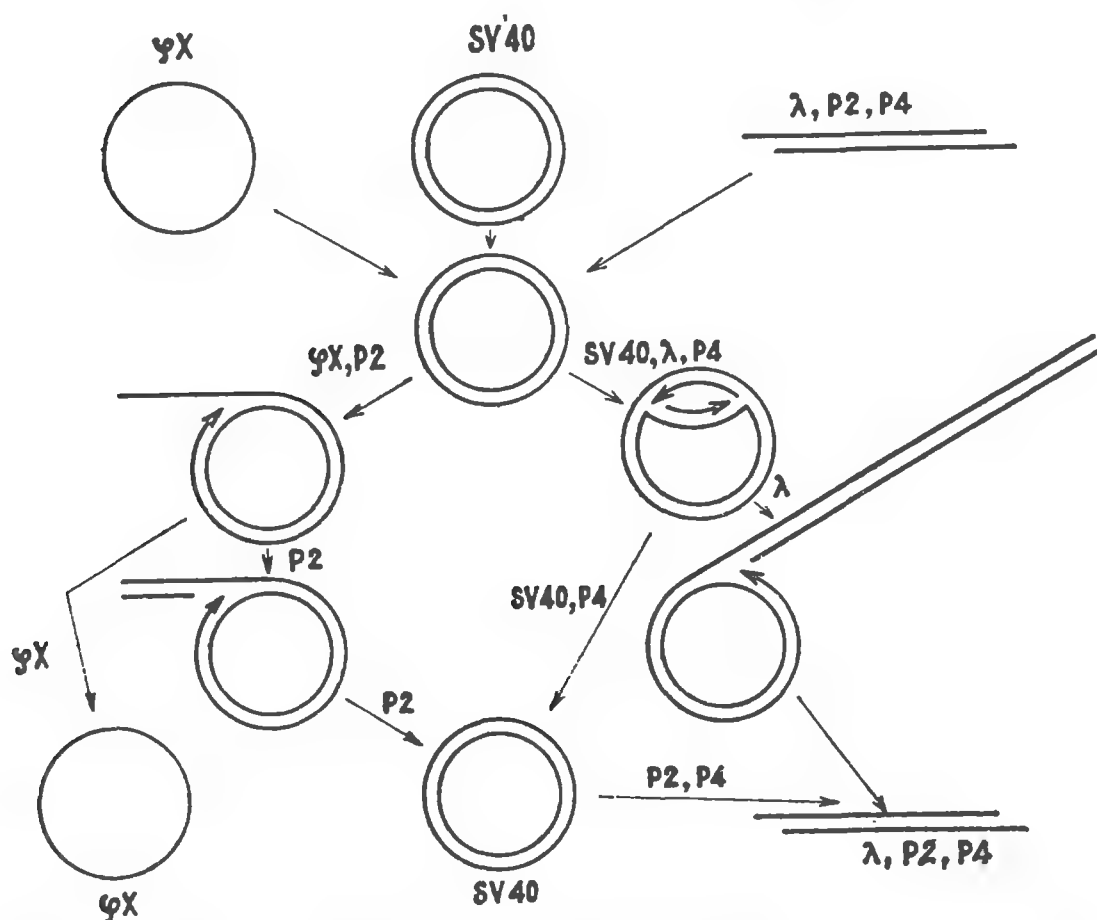


Рис. 148. Схема, иллюстрирующая блочный принцип организации систем репликации ДНК-геномов

ции ДНК-генома и возможность комбинирования этих блоков в разных сочетаниях можно назвать *блочным принципом организации репликационных систем*.

Ферментативное обеспечение репликации

Репликация ДНК-геномов требует участия ряда ферментов и дополнительных белков (см. с. 54). Важное значение среди них имеют те, которые обеспечивают инициацию репликации двухнитевой молекулы (генома или репликативной формы), т. е. процесс, который определяет выбор той или иной схемы репликации. Так, для систем с участием терминального (концевого) белка такую роль играют сам этот белок и фермент, осуществляющий его нуклеотидилирование. Системы с внутренней инициацией на двухнитевой матрице предполагают участие белков, узнающих *ori*-сайт. Во всех изученных случаях эти белки закодированы в вирусном геноме. Однако этими ключевыми белками отнюдь не исчерпывается список необходимых участников процесса репликации. Разумеется, обязательно наличие ДНК-полимеразы, а также — в зависимости от схемы репликации и конкретно вирусной системы — набора дополнительных ферментов и белков: праймаз, хеликаз, лигаз, ДНК-связывающих белков и ряда других. В разных вирусных системах эти белки могут иметь либо клеточную природу, либо могут быть закодированы в вирусном геноме. Как крайние примеры можно привести фаги фХ174 и Т4. В первом случае все многочисленные компоненты репликационного аппарата, за исключением белка А, — это клеточные белки, а во втором — подавляющее их большинство имеет вирусное происхождение. Встречаются случаи, когда фермент, принимающий участие в репликации вирусного генома, — это как бы гибридная молекула, построенная из клеточной и вирус-специфической полипептидных цепей. Так, ДНК-полимераза фага Т7 состоит из двух субъединиц — одна из них закодирована в вирусном геноме, а другая представляет собой клеточный белок тиоредоксин.

Говоря о вирус-специфических репликационных белках, следует подчеркнуть, что во многих случаях в незараженной клетке имеются белки с аналогичной функцией. Причем в искусственных бесклеточных системах можно наблюдать, как клеточные ферменты «работают» на вирусной ДНК, а вирусные ферменты — на клеточной ДНК. Однако *in vivo* в зараженной вирусом клетке ситуация иная. Так, если вирус кодирует собственную ДНК-полимеразу (например, фаг Т4 или вирус герпеса), то репродукцию такого вируса может обеспечить только вирусный фермент и этот вирусный фермент не катализирует синтез клеточной ДНК в зараженной клетке. Это кажущееся противоречие — высокая специфичность по отношению к матрице *in vivo* и низкая *in vitro* — имеет

следующее объяснение. Белковые компоненты репликационного аппарата обычно функционируют не изолированно, а в виде специфических межмолекулярных комплексов, которые формируются благодаря белок-белковым взаимодействиям. Поэтому замена некоего вирусного белка на клеточный с аналогичной функцией может оказаться столь же бесполезной, как замена в часах шестеренки аналогичной деталью, взятой из механизма другого размера. Впрочем, частичная подмена функций некоторых вирус-специфических репликационных белков (например, таких, как лигазы или хеликазы) соответствующими клеточными белками может иногда оказаться возможной.

Таким образом, набор ферментативных активностей, необходимых для репликации генома данного вируса, задается схемой репликации, а происхождение соответствующих ферментов — вирусное или клеточное — может быть разным. Вирусы с крупным геномом обычно имеют собственный набор репликационных ферментов и дополнительных белков, что, очевидно, делает вирусную репликационную систему более автономной от клетки-хозяина.

Интеграция вирусных и клеточных геномов

Помимо разобранный автономной репликации вирусных ДНК-геномов в ряде случаев — прежде всего у умеренных вирусов — реализуется существенно иной способ воспроизведения вирусной ДНК. Речь идет об *интеграции вирусных и клеточных геномов*. Такая интеграция может осуществляться несколькими способами.

а. Интеграция генома фага лямбда (сайт-специфическая рекомбинация)

С некоторой вероятностью, зависящей как от внешних условий, так и от генетических особенностей вируса и клетки, ДНК фага λ ковалентно встраивается в хромосому *E. coli*. Поскольку в рекомбинацию в данном случае вовлечены, как правило, совершенно определенные участки фагового и бактериального геномов (*att P* и *att B* соответственно), принято говорить, что имеет место *сайт-специфическая рекомбинация*. Общая схема этого процесса представлена на рис. 149. Участки *att P* и *att B* имеют одинаковую «сердцевидную» нуклеотидную последовательность (ее обозначают O) и различающиеся «плечи» (P и P', B и B'). Соответственно *att P* = POP', а *att B* = BOB'. Для осуществления интеграции необходимы два белка: *интеграза*, закодированная в фаговом гене *int*, и клеточный белок IHF (англ. integration host factor — клеточный фактор интеграции). Роль этих белков двоякая. Во-первых, они

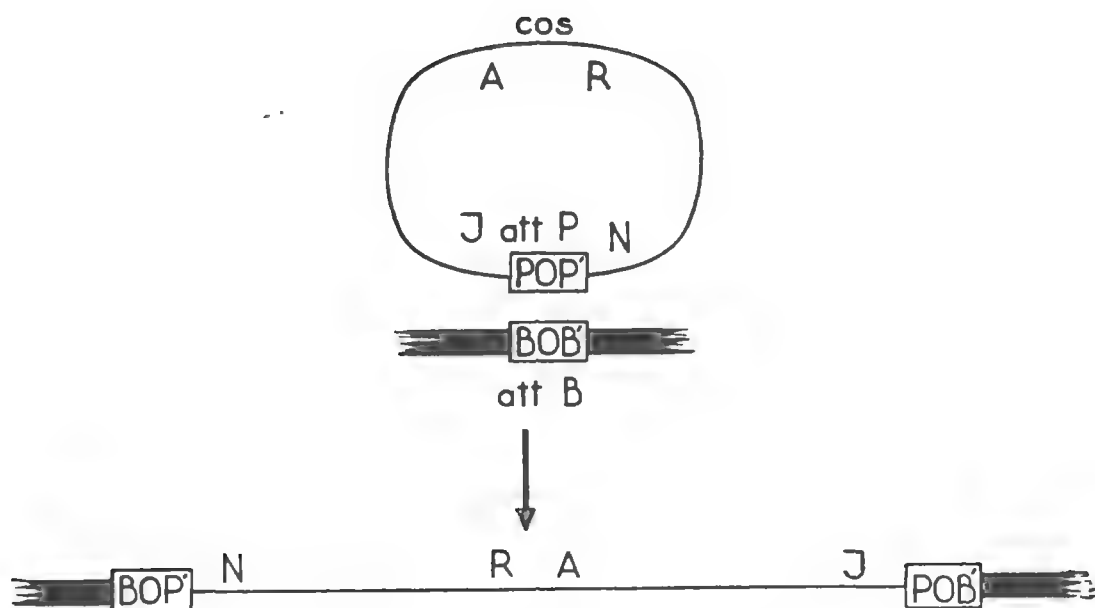


Рис. 149. Общая схема интеграции генома фага λ и клеточной ДНК
Тонкие линии — фаговая ДНК; жирные — ограниченные участки клеточной ДНК

специфически присоединяются к участкам *att P* и *att B*, переводят *att P* в более компактную структуру (так называемую *интасому*) и способствуют взаимному сближению *att P* с *att B*. Во-вторых, фаговая интеграза катализирует собственно рекомбинационный процесс. Предположительный механизм рекомбинации (рис. 150) основан на том, что интеграза проявляет свойства топоизомеразы I. Соответственно она может вносить временные однонитевые разрывы в ДНК-дуплекс. Полагают, что в месте взаимодействия *att P* и *att B* возникает как бы четырехнитевая структура ДНК. Внесение интегразой/топоизомеразой разрывов в соответствующие места фагового и клеточного геномов и последующее воссоединение концов в определенной комбинации приводят к интеграции этих геномов. Интегрированный вирусный геном называют *провирусом* или — в случае фагов — *профагом*. Отметим, что порядок расположения генов в профаге λ отличается от такового в зрелой молекуле фаговой ДНК (см. рис. 149 и 154); это объясняется тем, что участки *cos* и *att P* не совпадают. Профаг имеет совершенно определенную (единственную) ориентацию относительно клеточных генов, между которыми он находится.

Выщепление профага из клеточной хромосомы — это также результат сайт-специфической рекомбинации, но на этот раз между участками *BOB'* и *POB'*, которые находятся на концах интегрированной вирусной ДНК (рис. 149). Особенность реакции выщепления — потребность в дополнительном белке, продукте фагового гена *xis*. Такая потребность возникает вследствие того, что фаговая интеграза сама по себе слабо взаимодействует с участком *POB'*.

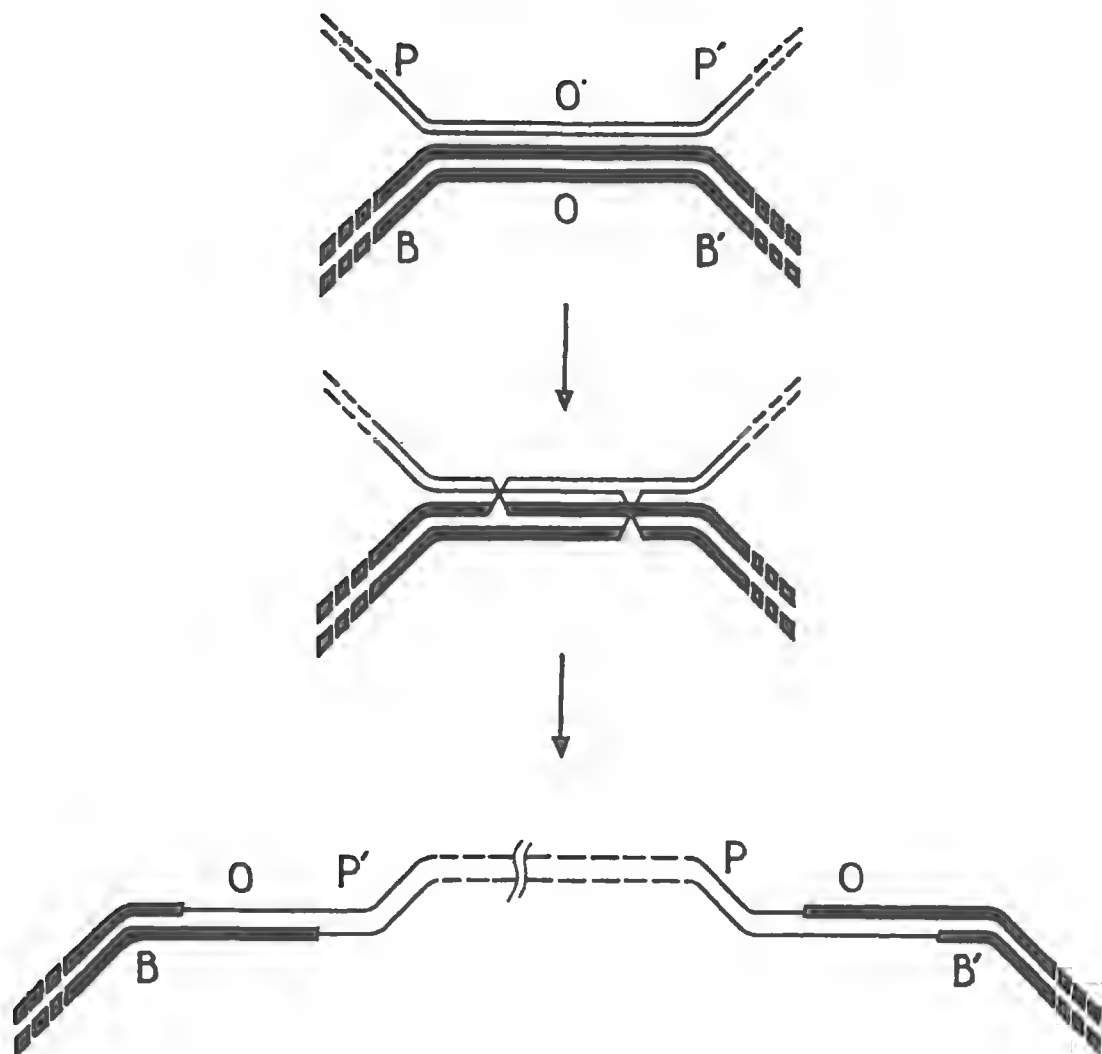


Рис. 150. Механизм сайт-специфической рекомбинации, ведущей к интеграции ДНК фага λ и клеточной ДНК

Тонкие линии — цепи фаговой ДНК; жирные — цепи клеточной ДНК; изображены лишь небольшие участки кольцевых геномов фага и бактерии

б. Интеграция и репликация генома фага μ (репликативная транспозиция)

Геном фага μ — линейная двухнитевая молекула ДНК, содержащая свыше 30 т. п. н. — имеет важную особенность: как левый (L), так и правый (R) концы молекулы содержат клеточные, а не вирус-специфические нуклеотидные последовательности. В определенном смысле можно сказать, что геном фага μ всегда находится в форме профага, поскольку он заключен между последовательностями клеточных ДНК. У индивидуальных молекул вирусной ДНК L-концы разные и R-концы разные, даже если популяция геномов образовалась в одной клетке. На границе L- и R-концов с вирус-специфической ДНК обнаруживается прямой повтор клеточной ДНК длиной в 5 п. н.; в разных молекулах вирусной ДНК

этот повтор имеет различную первичную структуру. Эти «странности» прямо связаны со способом репликации фага Ми.

После попадания фаговой ДНК в клетку происходит циркуляризация этой молекулы, однако образующиеся кольца (в отличие от ДНК фага λ) не имеют ковалентно-непрерывную структуру: просто концы молекулы сближены, и это сближение осуществляется при участии фагоспецифических белков.

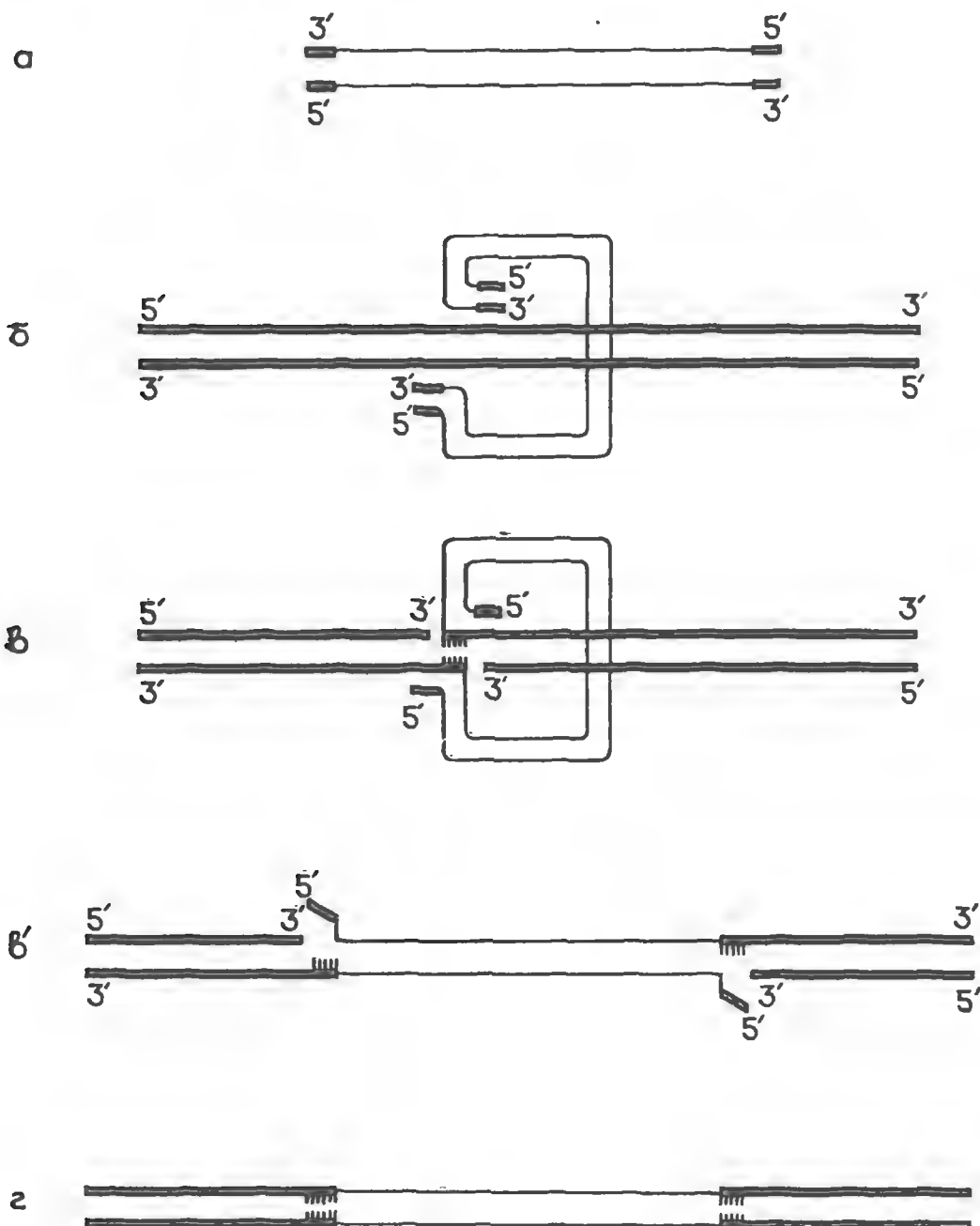


Рис. 151. Схема интеграции генома фага Ми и клеточной ДНК:

а — родительская фаговая ДНК с клеточными последовательностями на концах; б — сближение фаговой и клеточной ДНК; в — внесение одноцепочечных разрывов в клеточную и фаговую ДНК и ковалентное соединение 5'-концов цепей клеточной ДНК с 3'-концами цепей фаговой ДНК (г — та же структура, но в развернутом виде); д — репарация брешей и ковалентное соединение 3'-концов цепей клеточной ДНК с 5'-концами цепей фаговой ДНК; тонкие линии — цепи фаговой ДНК; жирные — цепи клеточной ДНК

Последующие события в схематическом виде представляются следующим образом (рис. 151). Участок фаговой ДНК со сближенными концами контактирует с каким-либо участком клеточной хромосомы, причем это может быть любой (или почти любой) участок клеточной ДНК. Далее под действием вирус-специфических белков происходит рекомбинация. В обе цепи клеточной ДНК на расстоянии пяти нуклеотидов вносятся одностебельные разрывы; кроме того, одностебельные разрывы вносятся в вирусную ДНК — по границе между L- и R-концами и вирус-специфическими последовательностями. При этом выступающие 5'-концы клеточной ДНК ковалентно соединяются с 3'-концами вирус-специфической ДНК. «Старые» L- и R-концы фаговой ДНК удаляются, и после репарации бреши фаговый геном оказывается встроенным в клеточную хромосому и окруженным вновь появившимся повтором клеточной ДНК длиной 5 п. н. Возможны две разные ориентации профага относительно клеточных генов; расположение генов в профаге и в ДНК вирусной частицы одинаково.

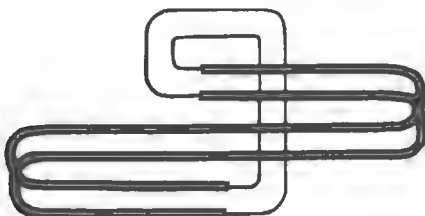
Необычной особенностью репликации ДНК фага М₁ является то, что, во-первых, все вновь синтезированные копии фагового генома оказываются в состоянии профага (т. е. включены в клеточную хромосому) и, во-вторых, фагоспецифическая последовательность нуклеотидов, которая послужила матрицей для образования дочерних геномов, остается в клеточной хромосоме на том же месте, где она находилась до репликации. Другими словами, репликация идет без выщепления «резидентного» профага и, по существу, представляет собой *репликативную транспозицию*. Вероятная схема этого процесса представлена на рис. 152. Фагоспецифические белки обеспечивают сближение концов профага, интегрированного в клеточную хромосому (аналогично тому, как они это делают с проникшей в клетку молекулой ДНК фага). Участок хромосомы, в котором сближены концы профага, контактирует с другим участком этой же хромосомы или с какой-либо другой находящейся в клетке молекулой ДНК. В этом «свежем» участке появляется ступенчатый разрыв (два одностебельных разрыва на расстоянии 5 п. н.); возникают одностебельные разрывы и по обеим границам резидентного профага. Выступающие 5'-концы клеточной ДНК соединяются с 3'-концами вирус-специфических последовательностей, а 3'-концы клеточной ДНК выполняют роль затравки. Таким образом, инициация раунда репликации представляет собой в этом случае вариант рекомбинационной инициации. В результате полуконсервативной репликации и последующих процессов репарации в клеточной хромосоме оказывается две копии профага; в каждой из них одна из цепей происходит из резидентного профага, а вторая синтезирована заново. При повторении этого процесса количество профагов в клеточной хромосоме может достигать сотни.

Образование зрелых геномов фага М₁ сопряжено с формиро-

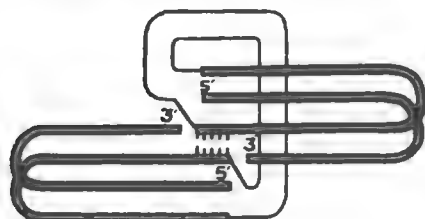
a



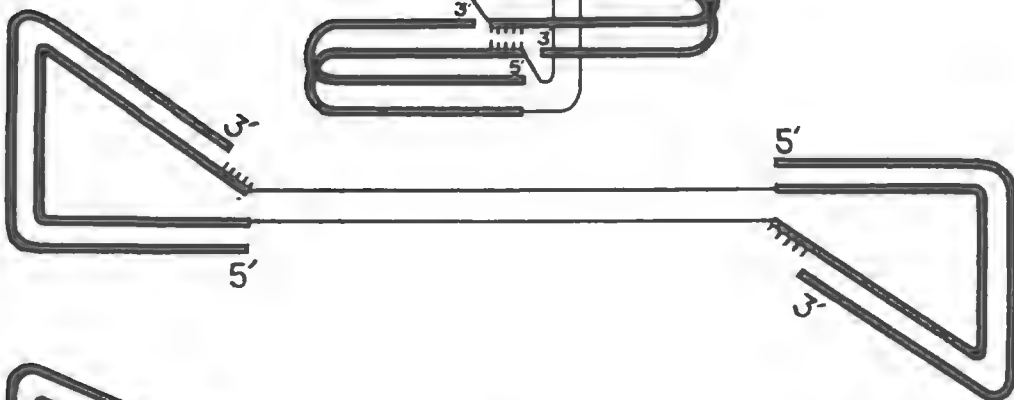
b



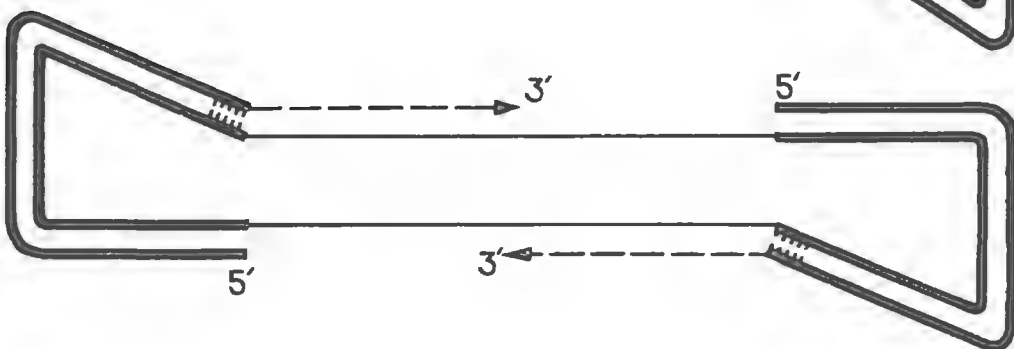
c



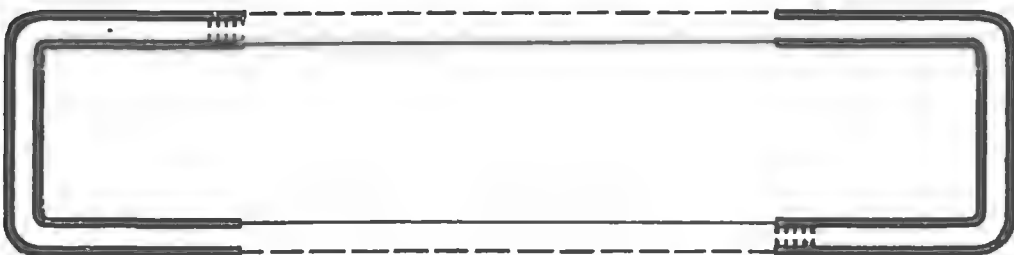
d



e



f



ванием вирусных частиц. При накоплении в клетке предшественников фаговых головок в клеточную ДНК на определенном расстоянии от левого конца профага вносится двухнитевой разрыв. С этого места — с образовавшегося L-конца — начинается упаковка фаговой ДНК в головку. Упаковывается столько ДНК, сколько может поместиться в головку, а далее вносится второй двухнитевой разрыв. Поскольку вместимость головки превышает объем ДНК, соответствующий вирус-специфической последовательности, этот второй разрыв приходится также на клеточную ДНК: возникает R-конец зрелой молекулы фаговой ДНК.

в. Другие виды интеграции

Для умеренных вирусов состояние провируса либо обязательная форма существования вирусного генома (например, у фага Mu), либо одна из форм хранения и воспроизведения генетической информации. В этих случаях выработаны специальные системы, обеспечивающие эффективную и правильную интеграцию вирусного и клеточного геномов. Однако включение вирусного генома или его фрагментов в клеточную хромосому не так уж редко происходит и в клетках, зараженных другими ДНК-содержащими вирусами. Такое включение достигается обычно за счет активности клеточных рекомбинационных систем и, как правило, является тупиковым путем: полноценный вирусный геном из таких интегрированных последовательностей чаще всего выщепиться не может. Другое дело, что для клетки приобретение вирус-специфических последовательностей может оказаться не безразличным. Если такая клетка выживет, то ее биологические свойства могут измениться; в частности, она может приобрести фенотип опухолевой клетки.

2. ТРАНСКРИПЦИЯ ГЕНОМА ДНК-СОДЕРЖАЩИХ ВИРУСОВ (ДНК→РНК)

Молекулярный механизм транскрипции — синтез молекул РНК на двухнитевой ДНК-матрице при помощи ДНК-зависимых РНК-полимераз — принципиально сходен в вирусных и клеточных

Рис. 152. Схема репликативной транспозиции профага Mu:

а — клеточная хромосома, содержащая профаг; *б* — сближение концов резидентного профага с новым участком клеточной ДНК; *в* — внесение одноцепочечных разрывов в этот участок клеточной ДНК и по границе между клеточной ДНК и резидентным профагом; ковалентное соединение 5'-концов цепей клеточной ДНК с 3'-концами цепей фаговой ДНК (*в'* — та же структура, но в развернутом виде); *г* — синтез комплементарных цепей профага; *д* — продукт репликативной транспозиции — клеточная хромосома, содержащая дополнительную копию профага; тонкие линии — цепи ДНК профага (сплошные — родительские, прерывистые — вновь синтезированные); жирные линии — цепи клеточной ДНК

системах. Особый интерес к транскрипции вирусных геномов обусловлен разнообразием способов регуляции этого процесса и относительной легкостью их изучения.

Необходимость регуляции транскрипции в вирусных системах связана с тем, что потребность в разных вирус-специфических белках разная: структурные белки требуются, как правило, в больших количествах, чем белки-ферменты. Кроме того, на ранних стадиях инфекции нужны преимущественно белки, обеспечивающие репликацию генома, а на поздних — белки, принимающие участие в формировании вирионов. Поэтому биологически целесообразно, чтобы разные вирусные гены считывались с разной эффективностью и чтобы эта эффективность изменялась по ходу репродукции вируса.

Считывание или несчитывание данного участка ДНК-матрицы зависит от расположения и свойств регуляторных сигналов двух типов: тех, которые обеспечивают начало транскрипции (промоторы и др.), и тех, которые прекращают транскрипцию (терминаторы) (см. гл. VII). Активность сигналов обоих типов может изменяться во времени — вплоть до их появления на определенной стадии как бы заново или, наоборот, до полного исчезновения.

Образование функционально активных молекул РНК помимо собственно транскрипции нередко включает те или иные пост-транскрипционные изменения или, как говорят, процессинг первичных транскриптов.

Характер образующихся транскриптов и способ регуляции транскрипции сильно зависят от того, имеем ли мы дело с вирусом прокариот или вирусом эукариот. Дело в том, что в клетках прокариот возможна множественная внутренняя инициация трансляции на полицистронных матрицах, тогда как в эукариотных молекулах РНК обычно реализуется единственная точка инициации трансляции и эти молекулы, как правило, функционально моноцистронны.

Регуляция транскрипции ДНК-генома вирусов прокариот

а. Транскрипция генома нитчатых фагов

Транскрипция генома нитчатых фагов (M13, *fd* и др.), в котором закодировано 10 белков, осуществляется клеточной ДНК-зависимой РНК-полимеразой. Экспрессия разных генов происходит с разной интенсивностью, которая заметно не меняется по ходу инфекции (т. е. временная регуляция отсутствует). Фаговые гены могут быть разбиты на две группы. Первая захватывает активно транскрибируемые гены, в которых закодированы белки, требующиеся в значительных количествах, в частности основной струк-

турный белок (продукт гена VIII) и ДНК-связывающий белок (продукт гена V). Область содержит несколько «сильных» промоторов и ограничена «сильным» ρ -независимым терминатором после гена VIII. Соответственно при транскрипции этого района образуется набор молекул мРНК разной длины, но с одинаковыми 3'-концами. Чем длиннее молекула мРНК, тем большее число белков она кодирует. Чем ближе к 3'-концу молекулы мРНК закодирован белок, тем в большем числе молекул мРНК представлен соответствующий цистрон.

Вторая область кодирует минорные структурные белки вириона и некоторые белки, принимающие участие в морфогенезе вирусной частицы. Эти белки нужны в относительно небольших количествах. Соответственно даже два главных из имеющихся здесь промотора можно отнести к разряду «слабых». Разделяющий две области фагового генома межгенный район содержит ряд регуляторных сигналов (в том числе *ori* для синтеза как (—), так и (+)цепей ДНК) и не транскрибируется, так как включает несколько ρ -зависимых и ρ -независимых терминаторов.

На этом примере видно, каким образом в результате соответствующего расположения промоторов и терминаторов разной силы можно обеспечить разную степень экспрессии разных вирусных генов.

б. Транскрипция генома фага лямбда

Регуляция транскрипции генома фага λ существенно богаче. Этот вирус принадлежит к числу умеренных: он либо вызывает типичную продуктивную инфекцию, либо встраивает свой геном в клеточную хромосому (переходит в состояние профага), сохраняя при этом способность при некоторых условиях высвободиться из «плена». Система регуляции транскрипции фага λ обеспечивает, во-первых, разную степень экспрессии разных генов, во-вторых, избирательное усиление или ослабление «работы» генов по ходу фаговой репродукции и, в-третьих, перестройку транскрипционного аппарата, необходимую для перехода в состояние профага или, наоборот, для выхода из этого состояния.

Транскрипция ДНК фага λ , как и в предыдущем случае, осуществляется бактериальной ДНК-зависимой РНК-полимеразой. Однако здесь вирусный геном кодирует ряд белков, модифицирующих взаимодействие этого фермента с ДНК-матрицей. Схематически главные этапы регуляции транскрипции фага λ можно представить следующим образом.

На фаговой ДНК имеется несколько промоторов, которые исходно узнаются клеточной РНК-полимеразой (рис. 153, а). С этих промоторов — их можно называть «истинно-ранним» — и начинается транскрипция проникшего в клетку вирусного генома. В частности, с промотора P_L транскрипция идет влево (по отношению к

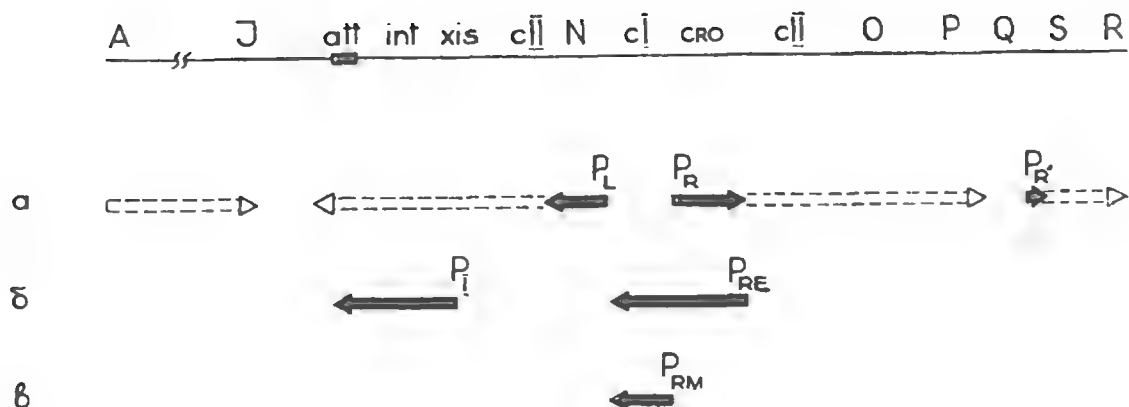


Рис. 153. Общая схема транскрипции генома фага λ

Сверху показано расположение некоторых генов на молекуле фаговой ДНК, а — промоторы, непосредственно узнаваемые клеточной РНК-полимеразой; черные стрелки — транскрипция в отсутствие фагоспецифических антитерминирующих факторов; пунктирные — транскрипция в присутствии белка N (с промоторов P_L и P_R) и Q (с промотора P_{R'}). Транскрипция генов A — J осуществляется с промотора P_R, так как эти гены в составе профага непосредственно примыкают к гену R (см. рис. 149); б — промоторы, активируемые белком CII; в — промоторы, активируемые белком CI

стандартной генетической карте) и заканчивается на р-зависимом терминаторе, расположенном левее гена N; с промотора P_R РНК-полимераза движется вправо (т. е. считывает другую цепь ДНК) и транскрибирует ген *cro*, вслед за которым также находится р-зависимый терминатор. На этой же стадии происходит abortивная транскрипция с промотора P_{R'}, но об этом процессе речь пойдет ниже (см. с. 296).

Белковые продукты генов N и *cro* — это регуляторы транскрипции. На следующей стадии репродукции фага особое значение имеет белок N — специфический антитерминатор. При наличии белка N транскрипция, начатая с промоторов P_L и P_R, продолжается соответственно влево от гена N и вправо от гена *cro* — РНК-полимераза игнорирует встречающиеся на ее пути терминирующие сигналы. Механизм антитерминирующего действия белка N до конца не расшифрован, но ясно, что он включает ряд компонентов: в частности, несколько клеточных белков (прежде всего РНК-полимеразу, а также полипептиды NusA и NusB) и определенные нуклеотидные последовательности *nut* (от англ. N utilization — использование белка N) в вирусной ДНК или РНК; белок N оказывает антитерминирующее действие только в том случае, если между соответствующим терминатором и промотором имеется последовательность *nut*, содержащая характерный несовершенный инвертированный повтор. Предполагают, что в момент, когда РНК-полимераза считывает область *nut*, она взаимодействует с белками N и Nus и в результате такого взаимодействия теряет способность «понимать» расположенные на ее дальнейшем пути определенные (прежде всего р-зависимые) терминирующие сигналы.

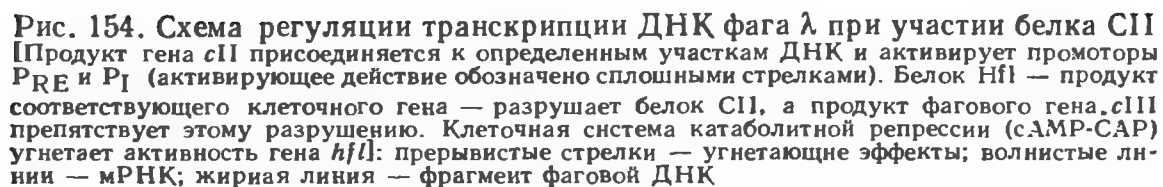
В результате антитерминирующего действия белка N существ-

венно возрастает число транскрибируемых фаговых генов — вновь «включенные» гены можно назвать «отсроченно-ранними». Интересно, что при этом не накапливается вирус-специфическая интеграза, несмотря на то что молекулы РНК-полимеразы, начавшие транскрипцию с промотора P_L и преодолевшие (благодаря белку N) ряд терминаторов, считывают ген *int*, в котором интеграза закодирована. Объяснение парадокса таково. Транскрипция, начатая с промотора P_L , не заканчивается на гене *int*, а продолжается дальше влево. Определенный участок образующегося транскрипта, расположенный за геном *int* (т. е. ближе к 3'-концу), комплементарен предшествующему участку этого же транскрипта; в результате возникает внутримолекулярная двухнитевая структура, которая разрушается клеточным ферментом РНКазой III. Это приводит к деградации мРНК для белка Int. Такой способ регуляции экспрессии генов, который осуществляется при участии нуклеотидных последовательностей, расположенных за геном, называют *ретрорегуляцией*.

Особо существенную роль в дальнейшем развитии событий играют белки — продукты фаговых генов *cII* и *cIII*. Первый из них способен присоединяться к определенным участкам ДНК фага λ . В результате здесь активируются новые промоторы для клеточной РНК-полимеразы, в частности промоторы P_{RE} и P_I (рис. 153, б), левонаправленная транскрипция с которых обеспечивает считывание соответственно генов *cI* и *int*. Продукты этих генов — репрессор *cI* и интегразы, — если они накапливаются в достаточных количествах, направляют судьбу зараженной клетки в сторону лизогении. В принятии такого важного решения участвует несколько «инстанций» — вирусных и клеточных регуляторных систем.

Прежде всего следует отметить, что белок *cII* внутри клетки нестабилен — он довольно быстро разрушается клеточной протеолитической системой, контролируемой генами *hfl* (англ. high frequency of lysogenization — высокая частота лизогенизации). Фагоспецифический белок *cIII* повышает устойчивость белка *cII* к протеолизу. С другой стороны, клеточная система катаболической репрессии, включающая сАМР и САР-белок (см. гл. VII), угнетает активность *hfl*-системы, подавляя, по-видимому, транскрипцию соответствующих генов (рис. 154). Благодаря этой довольно сложной регуляторной сети концентрация белка *cII* зависит от разнообразных внутренних и внешних условий. Так, например, при недостатке в среде глюкозы увеличивается внутриклеточная концентрация сАМР; это ведет к угнетению экспрессии генов *hfl* и, следовательно, к повышению концентрации белка *cII* и в конечном счете к лизогенизации.

Уже упоминалось, что взаимодействие белка *cII* с фаговой ДНК активирует новые промоторы, с которых считываются, в частности, гены *cI* и *int*. Транскрипция гена *int* с промотора P_I ведет к образованию функциональноактивной мРНК для интегразы. Дело в



Накопление интегразы способствует переходу фаговой ДНК в состояние профага, а появление репрессора CI подавляет активность почти всех фаговых генов. Последний эффект — результат присоединения репрессора CI к левой и правой операторным зонам (O_L и O_R). Каждая из этих зон включает по три близких по структуре несовершенных инвертированных повтора, разделенных короткими спейсерами (рис. 155). Лучше всего и прежде всего репрессор взаимодействует с самым правым участком (O_{R1}) правой зоны и самым левым участком (O_{L1}) левой зоны. Поскольку эти участки перекрываются с промоторами P_R и P_L , присоединение репрессора инактивирует промоторы.

294

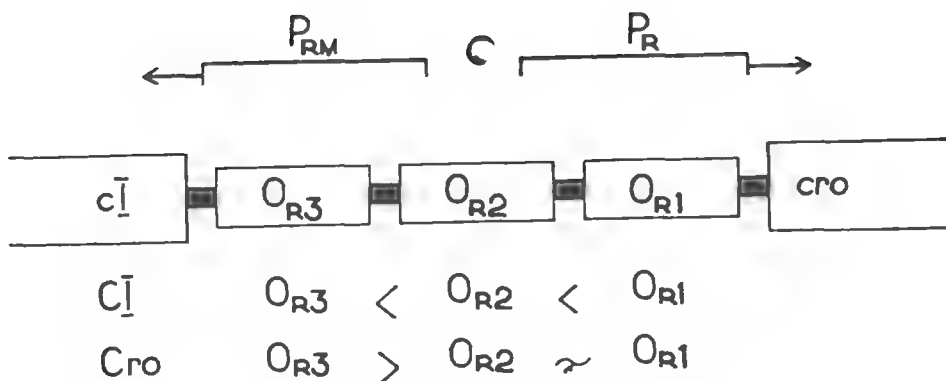


Рис. 155. Схема правой операторной зоны фага λ , расположенной между генами cI и cro

Снизу приведено относительное сродство репрессоров C_I и C_{cro} к трем операторным участкам (O_{R1} , O_{R2} и O_{R3}); сверху — расположение промоторов P_{R1} и P_{RM} , а также направление транскрипции

нирует, когда репрессор занимает участки O_{R1} и O_{R2} , а участок O_{R3} остается свободным. Активирование промотора в этом случае аналогично активированию промотора P_{RE} в результате присоединения белка CII к фаговой ДНК. Разумеется, в обоих этих случаях присоединение белка к ДНК — это условие, необходимое для создания промотора, но не достаточное. Необходима также определенная первичная структура ДНК по соседству с местом, к которому присоединен белок.

После активации промотора P_{RM} транскрипция гена cI автоматически поддерживается на постоянном уровне: при избыточном накоплении белка C_I он присоединяется к участку O_{R3} , что делает невозможной дальнейшую транскрипцию с промотора P_{RM} ; активность этого промотора восстанавливается, как только концентрация репрессора снижается до уровня, при котором его хватает только для присоединения к участкам O_{R1} и O_{R2} . Такой способ регуляции активности гена при помощи продукта этого же гена называют *аутогенной регуляцией*.

Репрессированное состояние профага может поддерживаться неопределенно долго при размножении лизогенных клеток. Однако при некоторых условиях (например, при активации клеточной SOS-системы; см. с. 79) репрессор разрушается (или инактивируется) и тогда происходит *индукция* профага. В результате инактивации репрессора C_I возобновляется транскрипция с промоторов P_R и P_L и синтезируются мРНК для белков C_{ro} и N . Белок N оказывает уже известное нам антитерминирующее действие, а белок C_{ro} обеспечивает переключение транскрипции на новые рельсы — на путь продуктивной инфекции. Белок C_{ro} , подобно белку C_I , — репрессор, взаимодействующий с правой (O_R) и левой (O_L) операторными зонами ДНК фага λ . Но сродство белка C_{ro} к разным операторным участкам иное, чем у белка C_I . В частности, в правой операторной зоне белок C_{ro} имеет наибольшее сродство к участку O_{R3} (рис. 155).

Присоединение белка $\text{C}_{\text{го}}$ к O_{R_3} предотвращает возможность левонаправленной транскрипции гена cI ; при этом правонаправленная транскрипция с промотора P_{R} разрешена (хотя и она будет ослабевать по мере увеличения концентрации белка $\text{C}_{\text{го}}$ и присоединения этого белка к участкам O_{R_1} и O_{R_2}); продолжается и левонаправленная транскрипция с промотора P_{L} .

Среди белков, образующихся на этой стадии, важную регуляторную роль играет продукт гена Q . Белок Q — это антитерминатор, обеспечивающий эффективную транскрипцию поздних фаговых генов с промотора P_{R} (см. рис. 153). Как уже упоминалось, клеточная РНК-полимераза узнает этот промотор, но в отсутствие белка Q происходит лишь абортивная транскрипция — два ρ -независимых мощных терминатора (аттенюатора), расположенные вблизи от промотора P_{R} , не оставляют возможности для образования полноценного транскрипта. Белок Q помогает РНК-полимеразе преодолеть эти (и только эти) терминаторы. Специфичность антитерминирующего действия белка Q задается последовательностью *qut*, локализованной между P_{R} и терминаторами (подобно тому, как последовательность *nut* задает специфичность антитерминирующего действия белка N). С промотора P_{R} транскрибируются *поздние* гены, в которых закодированы белки, необходимые для формирования фаговой частицы. На рис. 153 большинство поздних генов (А—J) располагаются на другом конце молекулы. Следует, однако, помнить, что на поздних стадиях инфекции транскрипция происходит на конкатемерных и кольцевых матрицах — именно поэтому поздние гены считываются с промотора P_{R} .

в. Транскрипция генома фага Т4

По ходу инфекции фагом Т4 происходит включение и выключение вирусных генов, как это было выяснено в пионерских работах Р. Б. Хесина и его сотрудников. Соответственно различают *истинно-ранние*, *отсроченно-ранние* (средние) и *поздние* гены и такие же промоторы. Структура истинно-ранних промоторов фага близка к таковой канонических промоторов клетки-хозяина. Это неудивительно, так как именно истинно-ранние промоторы узнаются клеточной РНК-полимеразой сразу же после попадания фаговой ДНК в клетку. Отсроченно-ранние промоторы фага Т4 имеют заметные отличия в области «—35». Активация этой группы промоторов через некоторое время после начала инфекции достигается, по-видимому, при помощи механизма, сходного с тем, который используется для активации промоторов P_{RE} и P_{I} белком CII фага λ : к соответствующим участкам фаговой ДНК присоединяется сначала фагоспецифический белок или белки (продукты гена или генов *mot*), а затем уже РНК-полимераза. Некоторые отсроченно-ранние гены фага Т4 считываются не только со «своих» промоторов, но также и с промо-

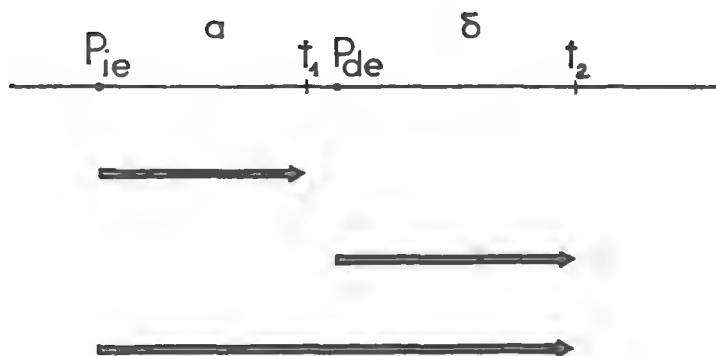


Рис. 156. Схема транскрипции ранних генов фага Т4

[Истинно-ранние гены (а) считываются только со «своих» промоторов (P_{ie}), а отсроченно-ранние (б) — либо со «своих» промоторов (P_{de}), либо при наличии антитерминирующих факторов — с промоторов истинно-ранних генов]: t_1 и t_2 — терминаторы

торов истинно-ранних генов, расположенных по соседству (рис. 156). Это становится возможным, поскольку вирус-специфический антитерминирующий фактор позволяет РНК-полимеразе преодолевать терминаторы, разделяющие соседние гены; при этом естественно образуются полицистронные мРНК, кодирующие как истинно-ранние, так и отсроченно-ранние белки.

Система регуляции транскрипции фага Т4 включает также элементы, которые еще не были описаны. В первую очередь следует назвать ковалентные и нековалентные модификации РНК-полимеразы, а также необходимость синтеза вирусной ДНК для включения поздних генов.

Продукты некоторых фаговых генов (*alt* и *mod*) способствуют введению АDP-рибозильных остатков в α -субъединицы РНК-полимеразы. Физиологическое значение такой модификации в точности не выяснено. На поздних стадиях инфекции наблюдается изменение полипептидного состава очищенных препаратов РНК-полимеразы: исчезает σ -субъединица (или ее содержание резко падает) и появляются вирус-специфические полипептиды, в частности продукты генов 55 (функциональный аналог σ -субъединицы), 33, а также некоторые другие. Полагают, что именно такая измененная, содержащая вирус-специфические субъединицы РНК-полимеразы способна узнавать поздние промоторы, структура которых отличается от структуры ранних промоторов как в районе «—35», так и в районе «—10». Но чтобы узнавание поздних промоторов было эффективным, ДНК-матрица должна находиться в «компетентном» состоянии. Молекулярная природа такого состояния не расшифрована, но ясно, что оно возникает при репликации фаговой ДНК.

Модифицированная, «поздняя», РНК-полимераза перестает узнавать промоторы ранних генов, что, очевидно, связано с утерей функционально-активной σ -субъединицы («узнающего» компонента хозяйского фермента).

Нельзя не упомянуть, что транскрипты некоторых фаговых генов (например, тимидилат-синтазы) подвергаются самосплайсингу (см. гл. VIII, раздел 2).

г. Транскрипция генома фага Т7

У фага Т7 встречается еще один способ временной регуляции транскрипции — образование фагоспецифической РНК-полимеразы. Схематически геном этого вируса можно разбить на две области: меньшая (~20 % генома) соответствует левому концу генетической карты и транскрибируется на ранней стадии инфекции, а большая содержит поздние гены. В начале ранней области располагаются по соседству несколько мощных промоторов, узнаваемых клеточной РНК-полимеразой, а заканчивается этот район не менее мощным р-независимым терминатором. Внутри ранней области имеются дополнительные слабые промоторы и слабые р-зависимые терминаторы, которые обеспечивают тонкую настройку работы ранних генов.

После попадания фаговой ДНК в клетку бактериальная РНК-полимераза считывает раннюю область. Среди образующихся транскриптов — гигантские молекулы полицистронных мРНК, кодирующие все ранние белки. Эти мРНК могут функционировать в виде единых молекул, но обычно подвергаются посттранскрипционному процессингу — расщепляются клеточной РНКазой III до моноцистронных матриц (расщепление обеспечивается тем, что межцистронные участки этих мРНК способны формировать элементы с внутримолекулярной вторичной структурой, узнаваемые РНКазой III).

В числе продуктов ранних генов — фагоспецифическая РНК-полимераза, закодированная в гене 1. Это относительно простой фермент, который в отличие от бактериальной РНК-полимеразы содержит всего одну полипептидную цепь ($M_r=107\,000$). Вирусный фермент узнает иной набор промоторов — поздние промоторы, которые имеют сходные между собой, но не идентичные первичные структуры. Поздние промоторы расположены преимущественно в поздней области фагового генома, но встречаются и в ранней, в частности они предшествуют участку *ori*, с которого начинается репликация вирусной ДНК. Поздние гены транскрибируются с разной эффективностью и в определенной последовательности. Не все механизмы этой регуляции расшифрованы, но некоторые из них достаточно понятны. В частности, в поздней области есть районы, которые организованы сходно с активно транскрибируемым районом генома нитчатых фагов (см. с. 290): такие участки имеют несколько промоторов и ограничены общим сильным терминатором. Отсюда считывается набор молекул мРНК разных размеров, но с одинаковыми 3'-концами. Чем ближе ген примыкает к терминатору, тем чаще он представлен в таком наборе мРНК. С другой стороны, есть участки ДНК, которые содержат общий промотор и несколько последовательно расположенных относительно слабых терминаторов, ко-

которые с некоторой вероятностью могут быть преодолены РНК-полимеразой. В наборе транскриптов таких участков, наоборот, в большей степени представлены гены, примыкающие к общему промотору.

Среди ранних фаговых белков есть и другие регуляторы транскрипции, помимо новой РНК-полимеразы. Особо важную роль среди них играет продукт фагового гена 2. Этот белок прочно соединяется с клеточной РНК-полимеразой. В результате фермент инактивируется, и тем самым прекращается транскрипция ранней области фагового генома.

д. Транскрипция генома фага N4

Особенность фага N4 — наличие вирус-специфической РНК-полимеразы в составе зрелых вирионов. Этот фермент — крупный белок ($M_r \sim 320\,000$) — продукт позднего фагового гена. Он попадает в заражаемую клетку вместе с вирусной ДНК и обеспечивает транскрипцию ранних генов. Среди продуктов ранних генов имеется вторая вирус-специфическая РНК-полимераза, построенная из трех относительно некрупных неидентичных полипептидных цепей. Эта вторая РНК-полимераза узнает промоторы средних генов. Поздние гены фага N4 транскрибируются, по-видимому, РНК-полимеразой *E. coli*.

Регуляция транскрипции ДНК-геномов вирусов эукариот

Основа регуляции транскрипции в случае ДНК-содержащих вирусов эукариот та же, что и у ДНК-содержащих фагов, — взаимное расположение и «сила» промоторов и терминаторов. Но в эукариотных системах встречаются новые регуляторные элементы, прежде всего энхансеры (см. гл. IX). Кроме того, образование «зрелых» молекул мРНК у ДНК-содержащих вирусов эукариот обычно связано с разнообразными посттранскрипционными изменениями (процессингом) первичных транскриптов. Это обстоятельство вносит важный вклад в регуляцию экспрессии генов.

а. Транскрипция генома обезьянного вируса 40 (SV40)

В геноме SV40 закодировано шесть белков. На ранней стадии образуются два белка, так называемые *большой* (Т) и *малый* (t) *антигены*. Затем — после начала репликации вирусной ДНК — идет синтез главным образом трех структурных белков вариона (VP1, VP2 и VP3). Функция и динамика синтеза шестого белка — *аглопротеина* — изучены недостаточно. Регуляция экспрессии вирусного генома осуществляется прежде всего на уровне транскрипции. Так, вскоре после инфекции в зараженных клетках начинают накапли-

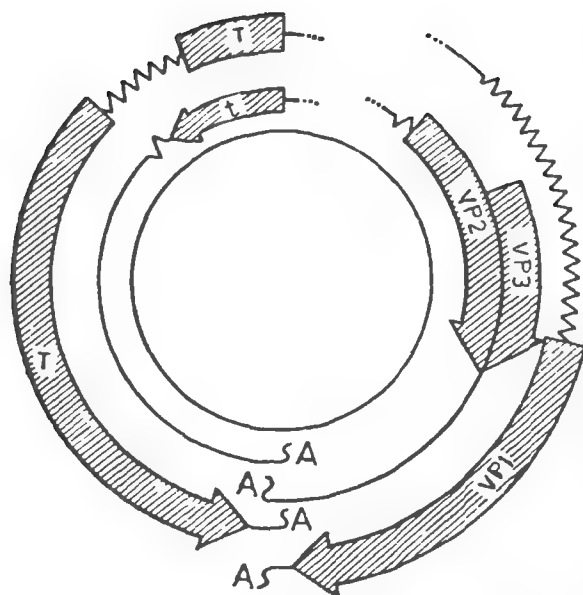


Рис. 157. Схема образования ранних и поздних мРНК SV40

Ранние и поздние первичные транскрипты считываются с разных цепей ДНК (в противоположных направлениях). «Зрелые» цитоплазматические мРНК образуются из первичных транскриптов в результате сплайсинга: внутренний круг — вирусный ДНК-геном, волнистая линия — поли(А). Заштрихованы открытые рамки считывания, зигзагообразные линии — сегменты первичных транскриптов, удаляемых при сплайсинге

ваться два класса ранних мРНК, кодирующих Т- и *t*-антигены, затем наблюдается преимущественный синтез двух классов поздних мРНК, кодирующих структурные белки; эти мРНК считываются с другой цепи вирусной ДНК и, следовательно, синтезируются в противоположном направлении по сравнению с ранними мРНК (рис. 157). Зрелые цитоплазматические мРНК — это продукты процессинга первичных транскриптов, которые в свою очередь представлены двумя классами: ранним и поздним. Рассмотрим сначала, как регулируется образование первичных транскриптов.

Синтез первичных транскриптов. Инициация синтеза обоих классов первичных транскриптов происходит в соседних участках генома — в области, которая насыщена регуляторными элементами

транскрипции, а также включает участок *ori* репликации ДНК (рис. 158). Оба класса транскриптов не имеют единственной, строго фиксированной стартовой точки; поэтому 5'-концы молекул мРНК внутри каждого класса несколько различаются по длине (особенно сильно такая микрогетерогенность выражена у поздних транскриптов). Тем не менее можно сказать, что промотор ранних мРНК имеет

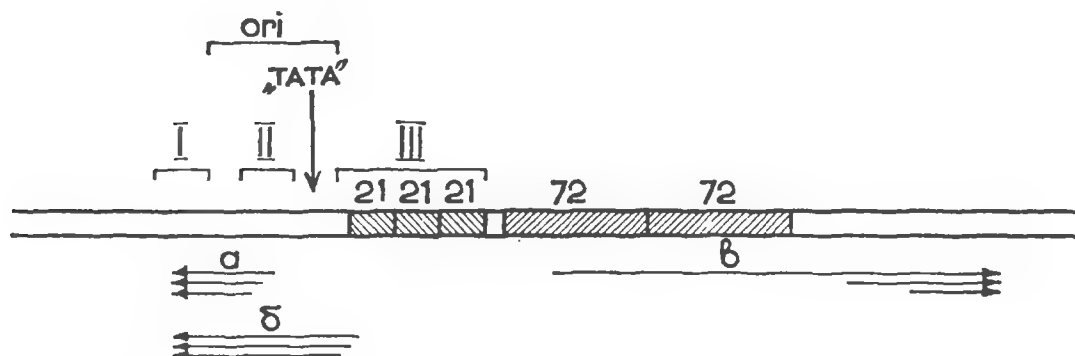


Рис. 158. Регуляторная область генома SV40

Заштрихованы повторы длиной 21 и 72 п. н.; римскими цифрами обозначены участки связывания Т-антигена; указано положение участка начала репликации (*ori*) и одного из регуляторных сигналов транскрипции (TATA); показаны также начальные участки транскриптов ранних генов на ранней (а) и поздней (б) стадиях инфекции и начальные участки транскриптов поздних генов (в); стрелки указывают направление транскрипции

типичный ТАТА-элемент (см. гл. X), расположенный примерно за 25 нуклеотидов до основной стартовой точки транскрипции. Кроме того, эффективность ранней транскрипции SV40 регулируется еще одним недостаточно охарактеризованным промоторным элементом, который находится внутри повторов длиной 21 п. н. Каждый из этих повторов содержит по два более коротких (6 п. н.) повтора, построенных из G-C-пар. Эти GC-богатые последовательности, по-видимому, взаимодействуют с клеточным белком — регулятором транскрипции. Нарушение взаимодействия приводит к значительному ослаблению синтеза ранних мРНК. Наконец, очень важным регулятором ранней транскрипции является энхансер, локализуемый внутри повтора, имеющего длину 72 п. н. Энхансер SV 40 проявляет все основные свойства, характерные для такого рода регуляторных элементов, и его активность также зависит от взаимодействия с клеточными регуляторными белками.

Все эти регуляторные элементы позволяют хозяйской РНК-полимеразе II осуществить эффективную транскрипцию ранних генов вскоре после попадания ДНК этого вируса в клеточное ядро. В результате процессинга ранних транскриптов (см. с. 302) образуются мРНК для ранних белков, а затем и сами эти белки. Один из них — Т-антиген — играет центральную роль в последующей перестройке транскрипции вирусного генома. Он вызывает ряд эффектов. Во-первых, взаимодействует с участками вирусной ДНК, связывающими Т-антиген (сильнее всего с участком I и слабее всего с участком III; см. рис. 158). В результате угнетается транскрипция ранних генов, в том числе и гена, кодирующего Т-антиген. Таким образом, Т-антиген проявляет здесь свойства репрессора, синтез которого подчиняется транскрипционной аутогенной регуляции. Впрочем, транскрипция ранних генов на поздней стадии прекращается не полностью. Она продолжается, хотя и со значительно меньшей эффективностью, но при этом стартовая точка транскрипции заметно смещается, так что ТАТА-элемент оказывается теперь внутри транскрибируемой последовательности (рис. 158). Механизм, обеспечивающий «позднюю» транскрипцию ранней области с новой стартовой точки, не расшифрован.

Второй эффект Т-антигена заключается в том, что он запускает репликацию вирусного генома и тем самым резко увеличивает число матриц («дозу гена»), с которых происходит считывание мРНК. Возможно, что увеличение числа молекул вирусной ДНК в клетке — одна из важных причин стимуляции транскрипции поздних генов. Дело в том, что на ранней стадии считывание этих генов угнетено в результате присоединения особого клеточного белка к повтору длиной 21 п. н. Количество этого белка-ингибитора в клетке весьма ограничено. Поэтому после начала репликации вирусного генома этот белок оказывается в дефиците и значительная доля молекул вирусной ДНК остается свободной от ингибитора.

Наконец, Т-антиген, вероятно, индуцирует образование клеточ-

ного белка, обладающего способностью, наоборот, специфически стимулировать позднюю транскрипцию.

Перечисленные эффекты скорее всего не исчерпывают многообразия регуляторных воздействий Т-антигена на транскрипцию генома SV40.

Не исключено, что в регуляции экспрессии поздних генов SV40 принимает участие и аттенуация транскрипции. РНК-полимераза II и на ранней стадии с некоторой эффективностью узнает поздний промотор, однако значительная часть образующихся при этом транскриптов обрывается (терминируется) после считывания ~90 нуклеотидов. Полагают, что в этой области имеется терминирующий сигнал, эффективность которого регулируется балансом терминирующих и антитерминирующих факторов, в число которых могут входить и вирус-специфические белки.

Процессинг первичных транскриптов. Процессинг первичных транскриптов SV40 включает те же реакции, что и процессинг большинства ядерных пре-мРНК эукариот: кэпирование 5'-конца, полиаденилирование 3'-конца и сплайсинг. Все эти реакции в случае SV40 осуществляются клеточными ферментами и при помощи тех же механизмов, которые реализуются в незараженной клетке. Остановимся здесь лишь на сплайсинге, точнее, на его значении для образования индивидуальных мРНК в связи с особенностями организации кодирующих последовательностей в вирусном геноме.

мРНК для Т-антигена и *t*-антигена — это разные молекулы, но образуются они из одного и того же класса ранних транскриптов. Обе мРНК имеют идентичные 5'-концы длиной чуть меньше 250 нуклеотидов и идентичные 3'-концы длиной несколько меньше 1700 нуклеотидов (рис. 157). Различие имеется лишь на стыке этих двух районов: при образовании мРНК для *t*-антигена в результате сплайсинга из первичного транскрипта удаляется 66 нуклеотидов, а при образовании мРНК для Т-антигена удаляется 346 нуклеотидов. В обеих мРНК инициаторный кодон предшествует участку сплайсинга, поэтому N-концевые последовательности Т- и *t*-антигенов идентичны. В мРНК для *t*-антигена терминирующий кодон также предшествует участку сплайсинга; имеющаяся открытая рамка считывания разрешает трансляцию относительно короткого полипептида. При образовании мРНК для Т-антигена этот терминирующий кодон удаляется в результате сплайсинга; возникает более протяженная открытая рамка считывания, и в состав Т-антигена входит уже примерно в четыре раза больше аминокислотных остатков.

Сплайсинг поздних транскриптов также дает начало двум группам поздних мРНК, одна из которых (19S) при трансляции образует полипептиды VP2 и VP3, а другая (16S) — VP1 и агнопротеин (рис. 157). Эти две группы мРНК различаются размерами сегмента, удаляемого из первичного транскрипта при сплайсинге. И хотя внутри каждой группы мРНК наблюдаются некоторые вариации в

длине удаляемого сегмента, в общем при образовании 16S мРНК он существенно длиннее, чем при образовании 19S мРНК. При этом в 19S мРНК отсутствует район, кодирующий агнопротеин, а протяженная открытая рамка считывания обеспечивает синтез VP2. В этой же рамке закодирован и VP3, но его синтез начинается на более удаленном от 5'-конца иницирующем кодоне (поэтому VP3 идентичен С-концевому фрагменту VP2). С другой стороны, в 16S мРНК сохранен ген агнопротеина, но отсутствуют оба инициаторных кодона, с которых в 19S мРНК начинается считывание VP2 и VP3. Синтез VP1 при трансляции 16S мРНК начинается на одном из последующих AUG-триплетов; при этом используется другая, еще более протяженная открытая рамка считывания. Как видно из рис. 157, начало гена VP1 перекрывается с концом генов VP2 и VP3.

б. Транскрипция генома аденовирусов

Геном аденовирусов содержит намного больше генов, чем геном SV40; регуляция экспрессии здесь сложнее и расшифрована не до конца. Рассмотрим только некоторые из элементов этой регуляции.

Транскрипция вирусного генома осуществляется в ядре клеточными РНК-полимеразами; при этом подавляющее большинство генов транскрибируется РНК-полимеразой II. Лишь два класса низкомолекулярных вирус-специфических РНК — VA1 и VA2 — синтезируются при помощи РНК-полимеразы III (эти РНК, по-видимому, принимают участие в регуляции трансляции вирусных матриц).

Принято разделять гены аденовирусов на ранние и поздние, хотя такое разделение, как и во многих других вирусных системах, несколько условно. К ранним относят гены, которые наиболее эффективно транскрибируются до начала репликации вирусной ДНК.

Имеется несколько районов, с которых считываются ранние мРНК, и один весьма протяженный район, в котором сосредоточены поздние гены (рис. 159). Каждый район кодирует несколько (обычно — много) разных мРНК. Причин этому несколько: во-первых, район может содержать не один промотор, и тогда синтезируется несколько разных первичных транскриптов; во-вторых, и это более важно, альтернативный посттранскрипционный процессинг приводит к образованию разнообразных молекул из идентичных первичных транскриптов.

Синтез первичных транскриптов. Вскоре после попадания вирусной ДНК в клеточное ядро начинается транскрипция района E1A (рис. 159). Этот район содержит регуляторные элементы, в частности ТАТА-последовательность и энхансер, обеспечивающие эффективное считывание при помощи РНК-полимеразы II. После процессинга первичных транскриптов (см. с. 305), выхода образовавшихся мРНК в цитоплазму и их трансляции появляется несколько

ранних белков. Эти белки транспортируются в ядро и оказывают там существенное влияние на последующую вирус-специфическую транскрипцию. Так, они регулируют эффективность считывания самого района E1A: основной эффект заключается в стимуляции экспрессии этого района; другими словами, эти ранние гены как бы «подстегивают» сами себя. Однако при определенных условиях продукты района E1A могут тормозить собственную работу. Природа этих сложных регуляторных взаимодействий еще не выяснена. Другой эффект экспрессии района E1A — стимуляция остальных ранних областей. Скорее всего стимуляция объясняется тем, что эти другие области (например, E2) имеют сходные с энхансерами регуляторные элементы, которые для проявления своей активности должны быть предварительно индуцированы. Роль соответствующих индукторов выполняют, по-видимому, белки, закодированные в районе E1A; одно из предположений заключается в том, что продукты гена E1A действуют не непосредственно на энхансер, а изменяют концентрацию клеточных регуляторов транскрипции. Имеют место взаимодействия и между другими ранними областями.

После начала репликации вирусной ДНК наблюдается ряд изменений в картине транскрипции. Работа ранних генов ослабляется; однако полного их выключения не происходит. Некоторые ранние гены на поздних стадиях считываются с иных промоторов, нежели в начале инфекционного цикла. Например, такая смена про-

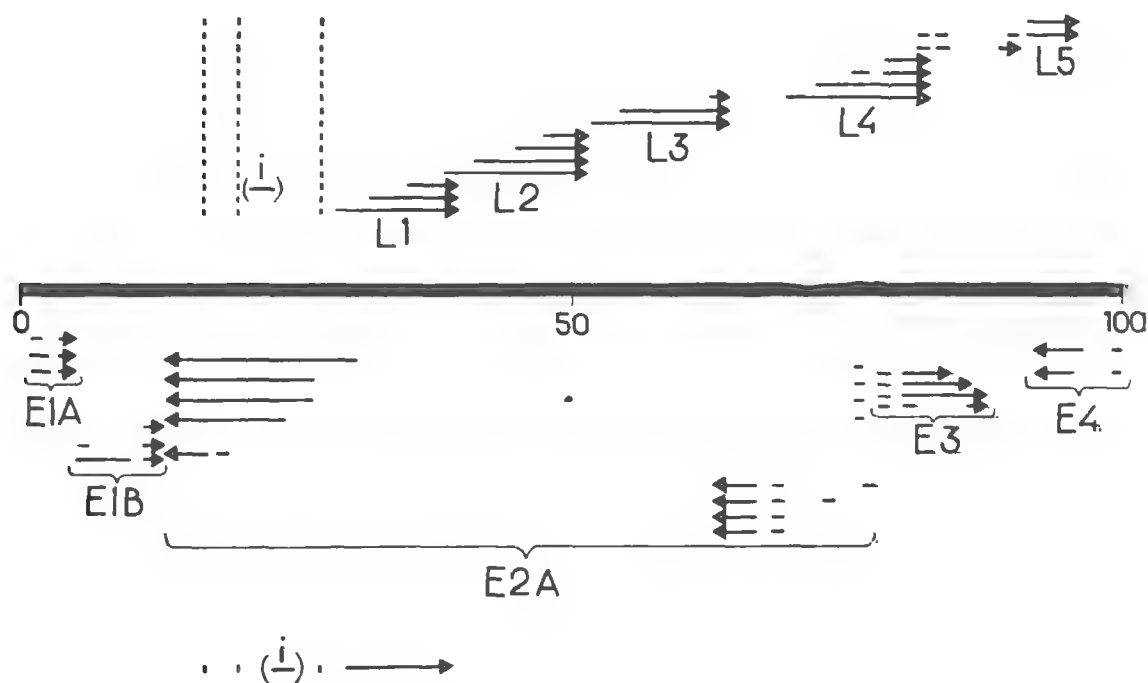


Рис. 159. Схема транскрипции генома аденовируса

Жирная сплошная линия — вирусный геном; наборы ранних мРНК (E1—E4) приведены под геномом, наборы поздних мРНК (L1—L5) — над геномом; показаны только экзоны; так, каждая поздняя мРНК содержит три (в одном случае — четыре) коротких «лидерных» экзона и один или несколько экзонов в «теле» молекулы; стрелки обозначают 3'-концы молекул мРНК

моторов происходит в области Е2, в которой закодирован ДНК-связывающий белок. Резко стимулируется в этот период считывание поздних генов. Следует отметить, что основной промотор поздних генов (расположенный на расстоянии примерно 16 % длины генома от «левого» его конца) относительно эффективно функционирует и на ранней стадии. Правда, в это время транскрибируется на весь поздний район, так как в его середине находится терминирующий сигнал. На поздней стадии, во-первых, усиливается транскрипция с того же самого промотора, а во-вторых, срабатывает антитерминирующий механизм и транскрипция теперь продолжается почти до «правого» конца молекулы аденовирусной ДНК (рис. 159). Механизмы этих двух регуляторных процессов до конца не расшифрованы.

Процессинг первичных транскриптов. Ядерные первичные транскрипты аденовирусного генома на своем пути до цитоплазматических зрелых мРНК претерпевают ряд превращений, и большинство этих реакций процессинга подвержено регуляции по ходу инфекции. Лишь «кэпирование» 5'-конца транскриптов, которое происходит еще до завершения их синтеза, осуществляется одинаково эффективно на разных стадиях цикла репродукции. Что касается полиаденилирования, то эта реакция вносит очень важный вклад в экспрессию аденовирусного генома. Особенно наглядно этот вклад виден при образовании поздних мРНК. В первичном транскрипте поздней области вирусного генома есть пять участков, несущих сигналы полиаденилирования (гексануклеотид AAUAAA и другие элементы). Полиаденилирование может произойти в любом из этих участков. Соответственно существует пять классов вирус-специфических РНК: L1—L5 (рис. 159); из данного первичного транскрипта может образоваться только одна молекула мРНК, принадлежащая к одному из этих пяти классов. От выбора того или иного участка полиаденилирования зависит относительная концентрация мРНК определенного класса. Образующийся на ранней стадии укороченный первичный транскрипт поздней области содержит три участка полиаденилирования, причем на этой стадии инфекции преимущественно используется самый первый из них (т. е. расположенный ближе к 5'-концу); в результате генерируется относительно много молекул мРНК класса L1. Однако на поздней стадии молекул класса L1 образуется меньше, чем, например, молекул класса L2.

Подавляющее большинство кэпированных и полиаденилированных транскриптов аденовирусного генома подвергается сплайсингу; этот процесс происходит в ядре и осуществляется в основном (если не исключительно) при помощи клеточных механизмов. Как правило, сплайсинг транскриптов данного класса может идти несколькими путями, т. е. в зрелой молекуле мРНК могут отсутствовать разные участки первичного транскрипта. Наглядный пример такого альтернативного сплайсинга дает опять процессинг поздних транскриптов (рис. 159).

Все многочисленные виды зрелых молекул поздних аденовирусных мРНК содержат так называемую лидерную область, которая состоит из трех, а иногда из четырех весьма коротких нуклеотидных последовательностей, происходящих из отдаленных друг от друга участков 5'-концевого сегмента поздних первичных транскриптов. Обычно эта область в белок не транслируется; исключение составляет лишь относительно редко встречающийся четырехкомпонентный лидер, включающий последовательность *i*. Кодированные районы расположены далее к 3'-концу — в «теле» мРНК. В разных молекулах мРНК, образовавшихся из транскриптов данного класса (например, L1), это «тело» имеет разную длину, причем, как правило, «ампутируются» 5'-концевые последовательности «тела». Такой способ генерации молекул мРНК представляется биологически весьма целесообразным, если учесть, что в клетках эукариот трансляция мРНК начинается преимущественно с первого от 5'-конца AUG-триплета. Поэтому в разных молекулах вирус-специфических мРНК данного класса роль иницирующего кодона выполняют разные AUG-триплеты, часто находящиеся в разных рамках считывания. Тем самым обеспечивается выражение компактно записанной генетической информации. Отметим, что при альтернативном сплайсинге говорить об интронах и экзонах можно только применительно к данной зрелой форме мРНК.

Важно, что по ходу инфекции эффективность того или иного варианта сплайсинга может изменяться.

Концентрация разных молекул вирус-специфических мРНК в цитоплазме регулируется еще на двух уровнях. Во-первых, при транспорте из ядра в цитоплазму. Уже говорилось, что на ранних стадиях образуется относительно мало транскриптов классов L2 и L3 (по сравнению с транскриптами класса L1), но все же они образуются. Однако в цитоплазму соответствующие мРНК почти не транспортируются, хотя на поздних стадиях специфических проблем с выходом из ядра мРНК классов L2 и L3 не наблюдается. Во-вторых, стабильность разных вирус-специфических мРНК в цитоплазме различна и может изменяться во времени. О механизмах регуляции на этих двух уровнях практически ничего не известно.

в. Транскрипция генома вируса осповакцины

ДНК вируса осповакцины — один из самых крупных вирусных геномов — содержит почти 200 т. п. н. и кодирует более сотни белков, синтез которых регулируется во времени. Важнейшая особенность транскрипционной системы этого вируса — ее локализация в цитоплазме зараженной клетки. Поэтому транскрипционный аппарат клетки, содержащийся в ядре, малодоступен для вируса,

который должен сам себя обеспечивать необходимыми ферментами и факторами. Все они (или по крайней мере большая их часть) закодированы в вирусном геноме; более того, они находятся в сердцевине вирусной частицы и вместе с ней попадают в заражаемую клетку. Это и понятно: иначе нельзя было бы начать транскрипцию вирусного генома.

Серцевина содержит, в частности, многие десятки молекул вирус-специфической ДНК-зависимой РНК-полимеразы. Этот крупный многокомпонентный фермент распознает промоторы ранних генов. Промоторы сильно обогащены А·Т-парами и удалены от стартовой точки транскрипции примерно на 30 п. н.; клеточные РНК-полимеразы эти промоторы не «узнают». Регуляторные элементы типа энхансеров в геноме вируса осповакцины не описаны.

Ранние гены вируса осповакцины обычно расположены группами; считываемые с них транскрипты иногда перекрываются. Характерна ситуация, когда транскрипты имеют идентичные 3'-концы и различаются по длине 5'-концевой последовательности. Очевидно, в таких случаях используются разные промоторы. Поскольку в клетках эукариот трансляция чаще всего начинается с первого (от 5'-конца) AUG-триплета, мРНК с разной длиной 5'-конца могут направлять синтез разных белков; при этом могут использоваться как одинаковые, так и неодинаковые открытые рамки считывания. Впрочем, не так уж редко ранние мРНК имеют идентичные 5'-концы, но различаются по длине 3'-концевого сегмента; в таких случаях речь, по-видимому, идет о недостаточно эффективной терминации транскрипции. Такие мРНК дают идентичные белковые продукты (или по крайней мере белки с идентичными N-концевыми последовательностями аминокислот).

После начала репликации вирусного генома транскрипция многих (хотя и не всех) ранних генов угнетается и активируется считывание поздних генов. Промоторы поздних и ранних генов различны; в частности, первые как будто несколько беднее А·Т-парами, чем вторые. Считают, что поздние промоторы узнаются модифицированной вирус-специфической РНК-полимеразой: модификация заключается в том, что в состав этого фермента включается субъединица клеточной РНК-полимеразы II.

Помимо смены промоторов на поздней стадии меняется и эффективность терминации. Поздние мРНК в среднем заметно длиннее ранних и весьма гетерогенны по длине. Частично эта гетерогенность связана с неэффективной терминацией, а частично обусловлена посттранскрипционными изменениями. Так, к 5'-концу транскрипта данного позднего гена могут быть присоединены нуклеотидные последовательности [в том числе и поли(А)], происходящие из транскриптов других (в том числе и весьма далеко расположенных) генов. Механизм этой реакции, напоминающей транс-сплайсинг (см. гл. VIII), неизвестен.

Кроме того, посттранскрипционная модификация вирус-специ-

фических РНК включает и уже известные реакции — кэпирование 5'-концов и полиаденилирование 3'-концов.

Посттранскрипционный процессинг осуществляется ферментами, закодированными в вирусном геноме.

* * *

При сравнении способов регуляции транскрипции у ДНК-содержащих бактериофагов и вирусов эукариот видно много общего: и там, и тут регуляция основана на изменении силы сигналов, обеспечивающих начало и окончание считывания данного транскрипта. И там, и тут транскрипционные единицы нередко охватывают несколько генов. Однако если у вирусов прокариот образующиеся при этом полицистронные мРНК могут давать при трансляции несколько белковых продуктов, то у вирусов эукариот для образования разных белков первичные транскрипты должны подвергнуться разным вариантам сплайсинга. Таким образом, полицистронность мРНК в прокариотных системах и альтернативный сплайсинг — в эукариотных могут выполнять сходные биологические функции: они позволяют объединять гены в группы, транскрибируемые с общего промотора и, следовательно, имеющие общий регуляторный элемент. В обоих случаях возможно объединение не только соседствующих генов, но также и перекрывающихся генов.

Помимо альтернативного сплайсинга вклад в образование индивидуальных мРНК из данного первичного транскрипта может вносить также и альтернативное полиаденилирование, как это было показано на примере поздних мРНК аденовирусов. Следовательно, есть основания для важного обобщения: процессинг первичных транскриптов у вирусов эукариот функционально связан с моноцистронностью эукариотных мРНК.

3. РЕПЛИКАЦИЯ/ТРАНСКРИПЦИЯ ВИРУСНЫХ ГЕНОМОВ, ВКЛЮЧАЮЩАЯ СИНТЕЗ ДНК НА МАТРИЦЕ РНК

Переходим к рассмотрению вирусных систем, в которых цикл репликации генома схематически можно разбить на две главные реакции: синтез РНК на матрице ДНК и синтез ДНК на матрице РНК. При этом в состав вирусной частицы в качестве генома может входить либо РНК (как у ретровирусов), либо ДНК (как у вирусов гепатита В и мозаики цветной капусты).

*Репликация/транскрипция генома
ретровирусов (РНК→ДНК→РНК)*

а. Структура генома

Ретровирусы — это обширная группа вирусов, представители которой различаются как по биологическим свойствам (например, по способности вызывать злокачественные опухоли), так и по морфоло-

тии; имеются и заметные вариации в структуре их геномов. Тем не менее эти геномы имеют следующие общие черты. 1. Вирусная РНК — одноцепочечная молекула длиной до десяти тысяч нуклеотидов — имеет полярность такую же, как у мРНК, т. е. представляет собой (+) нить; молекулы РНК экпированы на 5'-конце и полиадеенированы на 3'-конце. 2. В молекулах РНК есть прямой концевой повтор (*r*) длиной несколько десятков нуклеотидов (на 3'-конце повторяющаяся последовательность предшествует поли(А)-тракту). 3. В вирусной РНК записана информация для синтеза по крайней мере трех групп вирус-специфических белков: структурных белков сердцевин вириона (Gag-белков), ферментативных белков, принимающих участие в обратной транскрипции вирусного генома и в интеграции вирус-специфической ДНК и клеточной хромосомы (продуктов гена *pol*), и белков, входящих в состав наружной липопротеидной оболочки вириона (Env-белков). У некоторых ретровирусов есть дополнительные гены; нередко наблюдаются также всякого рода перестройки генома, что обычно ведет к дефектности вируса, т. е. к его неспособности размножаться без вируса-помощника. 4. Вирусная частица содержит две молекулы геномной РНК; таким образом, редкой (если не уникальной) особенностью ретровирусов является диплоидность их генома.

6. Обратная транскрипция

После взаимодействия вирусной частицы с чувствительной клеткой сердцевина вириона освобождается из липопротеидной оболочки и оказывается в цитоплазме. Сердцевина содержит помимо структурных белков вирусную РНК и фермент — обратную транскриптазу (ревертазу). Фермент полифункционален: он способен синтезировать ДНК, используя в качестве матрицы либо РНК, либо ДНК; кроме того, он обладает активностью РНКазы Н (т. е. разрушает цепь РНК, входящую в состав РНК/ДНК-дуплекса). О некоторых других активностях белков полимеразного комплекса будет сказано ниже. Подобно всем другим ДНК-полимеразам, ревертаза способна «работать» только при наличии затравки. При синтезе (—)цепи по матрице вирусной (+)РНК затравка необычна; ее роль выполняет одна из клеточных тРНК; для данного вируса — это вполне определенная тРНК (например, триптофановая, пролиновая и т. д.). Дело в том, что на расстоянии примерно 100—200 нуклеотидов от 5'-конца (для каждого вируса — это величина постоянная) имеется участок, который комплементарен 3'-концевой последовательности молекулы тРНК, используемой в данном случае в качестве затравки; этот участок часто обозначают как pbs (от англ. primer binding site — участок связывания затравки). Затравочная молекула тРНК находится в вирионе и попадает в клетку в виде комплекса с вирионной РНК.

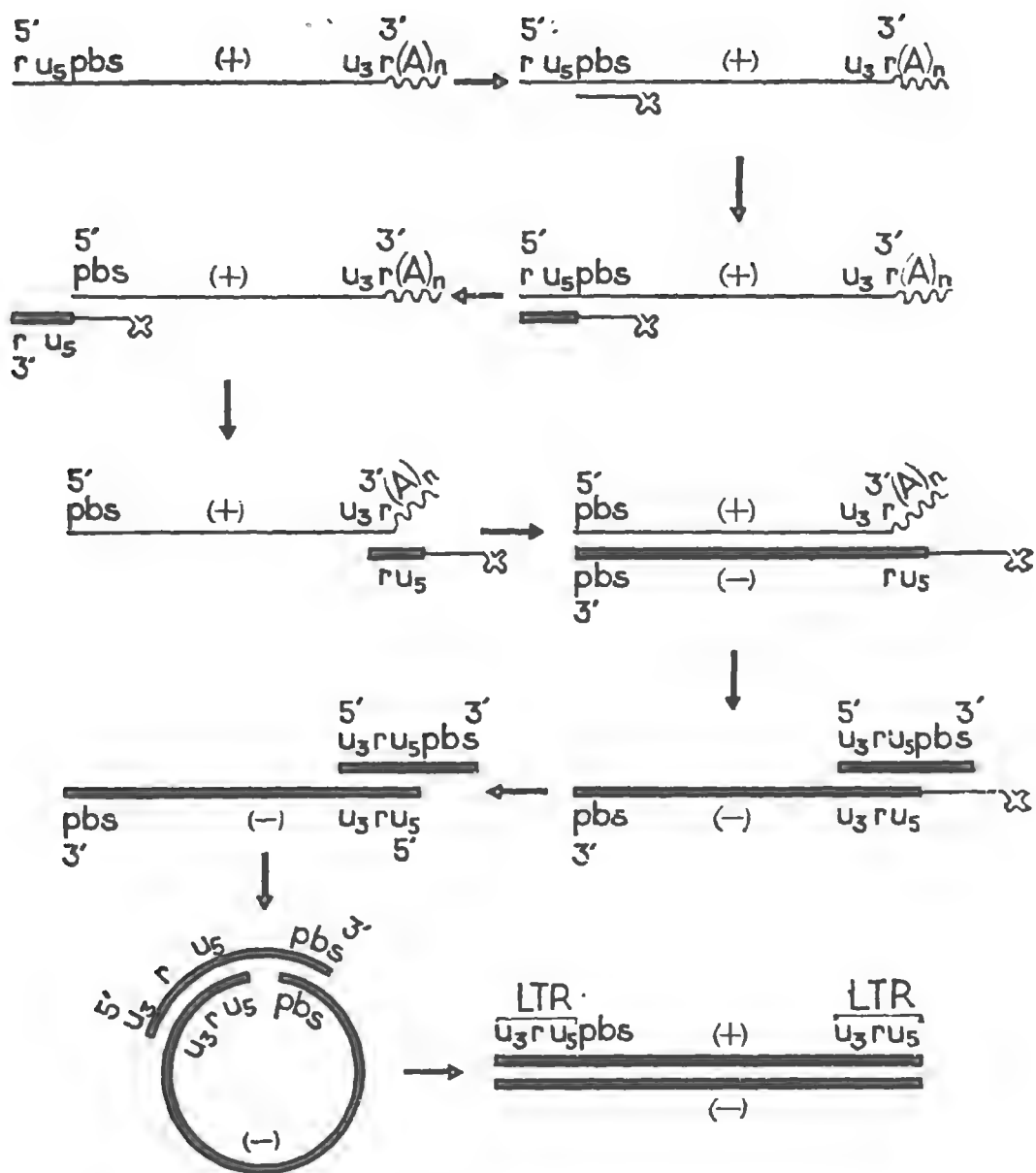


Рис. 160. Схема образования двухнитевой ретровирусной ДНК

Тонкие линии — цепи РНК [в том числе волнистая — поли(А), крестообразная структура — тРНК-затравка]; жирные линии — цепи ДНК; остальные обозначения см. текст

Первая стадия обратной транскрипции заключается в том, что, используя тРНК-затравку, ревертаза синтезирует сегмент ДНК, комплементарный 5'-концевой последовательности вирусной РНК (рис. 160). Этот сегмент принято называть (—)«strong-stop» ДНК, поскольку синтез ДНК после завершения копирования 5'-конца матрицы временно останавливается. (—)«Strong-stop» ДНК содержит последовательности, комплементарные концевому повтору *r* и примыкающему району, который называют районом *u5*, т. е. уникальным 5'-концевым районом.

На первый взгляд может показаться странным, что синтез ДНК начинается недалеко от 5'-конца матрицы — ведь при этом обра-

зуется только весьма короткий продукт. На самом деле такая тактика вполне понятна. Ведь (—)«strong-stop» ДНК содержит (в связи с наличием прямого концевой повтора) последовательность, комплементарную не только 5'-концу, но также и 3'-концу вирусной РНК, а как уже известно (см. раздел 1 этой главы), снятие ДНК-копии с 3'-конца матрицы — всегда проблема. У ретровирусов эта проблема решается весьма остроумно: комплемент 3'-конца матрицы производится в «удобном» месте, а затем переносится на «свое» место. Происходит это следующим образом. 5'-конец вирусной РНК, образующий дуплекс с (—)«strong-stop» ДНК, разрушается под влиянием присущей ревертазе активности РНКазы Н; в результате (—) «strong-stop» ДНК оказывается в однострессовой форме и поэтому может взаимодействовать с 3'-концом (с участком *r*) той же самой или другой молекулы вирусной РНК (рис. 160).

На следующей стадии в качестве затравки выступает уже (—) «strong-stop» ДНК; элонгация этой затравки приводит к синтезу (—) цепи ДНК, в которой, впрочем, может отсутствовать комплемент района *ru5*, поскольку соответствующий участок (+) матрицы был разрушен РНКазой Н (рис. 160).

Еще до завершения образования (—)цепи начинается синтез (+)цепи ДНК. В этом случае затравкой служит фрагмент (+)РНК, образовавшийся в результате действия РНКазы Н на вирус-специфический РНК/ДНК-дуплекс. Затравка образуется преимущественно в определенном месте, на расстоянии нескольких сотен нуклеотидов от 3'-конца вирусной (+)РНК. При использовании в качестве затравки этого фрагмента (+)РНК и в качестве матрицы уже синтезированного сегмента (—)цепи ДНК возникает участок вирус-специфической (+)ДНК, который называют (+)«strong-stop» ДНК. Этот участок соответствует определенному району вирусной (+)РНК — так называемому району *u3*, а также включает последовательность *ru5*. Кроме того, (+)«strong-stop» ДНК включает, по видимому, также последовательность *pbs*, для синтеза которой используется в качестве матрицы 3'-конец молекулы тРНК, послуживший затравкой при синтезе (—)«strong-stop» ДНК (рис. 160). Ситуация весьма необычна — вирус-специфическая ДНК образуется путем копирования клеточной тРНК! Затем молекула тРНК отщепляется, так как ее 3'-конец оказывается в составе РНК/ДНК-дуплекса и поэтому становится мишенью для РНКазы Н. Под действием этого же фермента удаляются и оставшиеся участки вирусной (+)РНК.

На следующем этапе происходит циркуляризация за счет взаимодействия между районом *pbs* в (+)«strong-stop» ДНК и комплементарной ему последовательностью в (—)цепи ДНК (рис. 160). Возникающая промежуточная форма удобна для завершения синтеза двухнитевой молекулы вирус-специфической ДНК. Действительно, 3'-конец (+)«strong-stop» ДНК теперь может быть использован в качестве затравки для завершения синтеза (+)цепи ДНК,

а недостающий участок (—)цепи ДНК (участок *ru5*) может быть скопирован с (+)«strong-stop» ДНК, которая в данной реакции выступает уже в роли матрицы (рис. 160).

Не все детали приведенной схемы образования вирус-специфической ДНК строго доказаны, и в дальнейшем в нее, возможно, будут внесены те или иные поправки. Тем не менее эта схема достаточно хорошо иллюстрирует общий принцип. Весьма важно, что схема объясняет одну чрезвычайно существенную особенность структуры вирус-специфических ДНК ретровирусов — молекулы вирусных ДНК длиннее молекул вирусных РНК, которые послужили матрицей для обратной транскрипции. Действительно, к 5'-концу (+)цепи вирусной ДНК добавилась последовательность *u3*, а к 3'-концу этой цепи — последовательность *u5*. В результате на концах молекулы вирус-специфической ДНК появился длинный (несколько сотен нуклеотидов) концевой повтор (ДКП, или LTR), имеющий структуру *u3ru5* (рис. 160).

Образование линейных молекул вирус-специфических ДНК осуществляется в цитоплазме зараженной клетки. Затем вирусная ДНК обнаруживается в ядре. Здесь линейные молекулы превращаются в кольцевые, причем циркуляризация происходит за счет взаимодействия между «тупыми» концами молекулы. Факторы, обеспечивающие переход линейных молекул в кольцевые, пока не изучены.

в. Интеграция

В настоящее время наиболее вероятной представляется такая последовательность событий, ведущих к включению вирус-специфической ДНК ретровирусов в клеточную хромосому (рис. 161). После образования кольцевой молекулы в месте стыка двух LTR возникает короткий несовершенный инвертированный повтор. Этот повтор выполняет функцию *att*, т. е. специфического участка интеграции. Участок *att* «узнается» вирус-специфическим ферментом, обладающим эндонуклеазной активностью — одним из продуктов гена *pol*, который попадает в клетку из заражающей вирусной частицы. Фермент вносит в обе цепочки молекулы вирус-специфической ДНК разрывы на расстоянии 4 нуклеотидов друг от друга. Этот же фермент вносит ступенчатый разрыв (на расстоянии 4—6 нуклеотидов) и в клеточную ДНК. Положение разрыва в клеточной ДНК не фиксировано. Далее происходит интеграция вирусной ДНК в хозяйскую хромосому. Предполагают, что механизм интеграции напоминает тот, который реализуется в фаговых системах, прежде всего у фага М₁ (см. раздел 1 этой главы), т. е. разрывы цепей ДНК и воссоединение гетерологичных нуклеотидных последовательностей осуществляет один и тот же фермент — особая топоизомераза (интеграза). Процессы типа репарационных (застраивание брешей и удаление одноцепочечных «хвостов») приводит к двум последствиям: во-

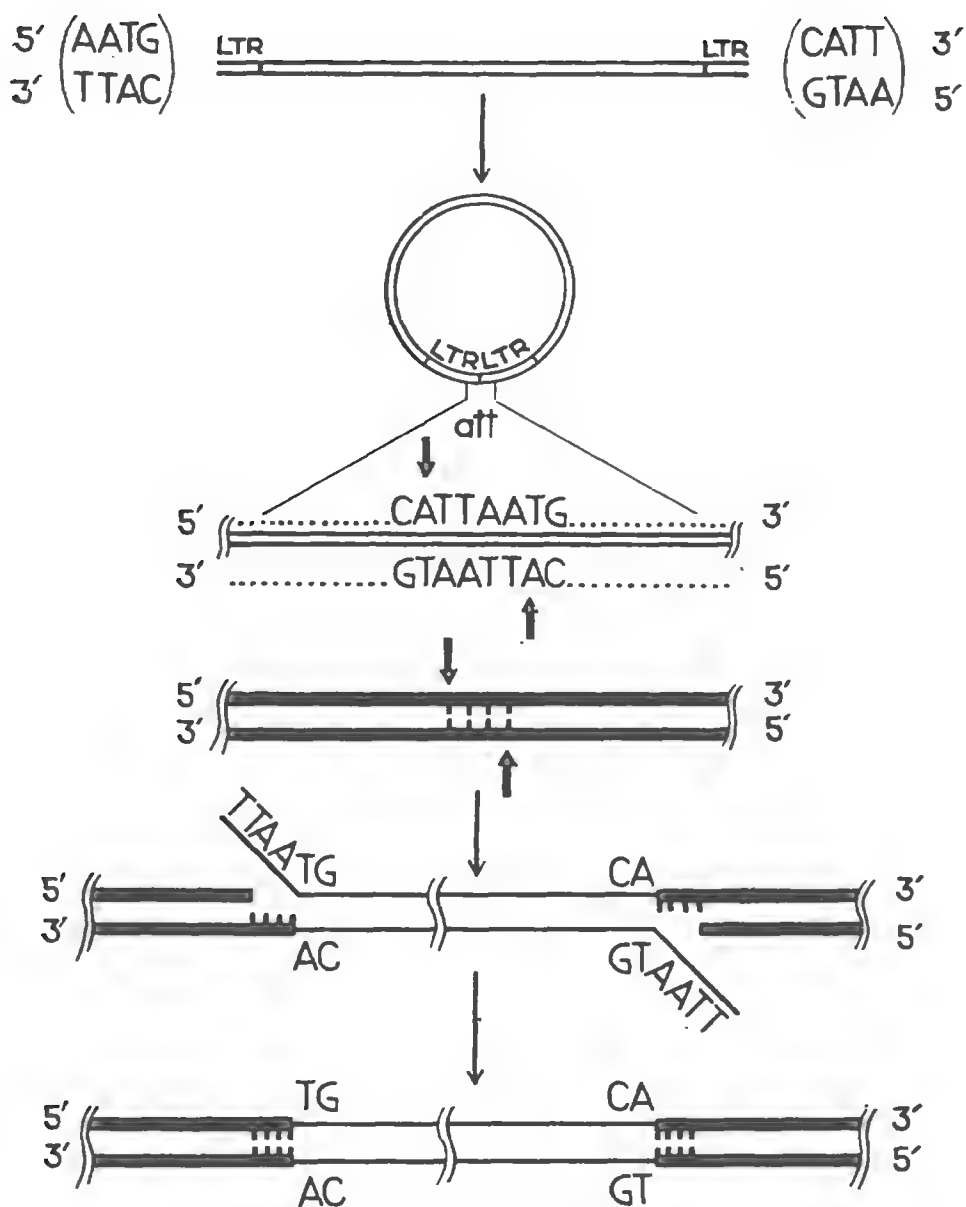


Рис. 161. Схема интеграции ретровирусной и клеточной ДНК

Тонкие линии — цепи вирус-специфической ДНК (сверху в скобках — концевые последовательности нуклеотидов); жирные линии — цепи клеточной ДНК; короткие стрелки — места разрывов, вносимых вирус-специфической эндонуклеазой

первых, в клеточной ДНК, фланкирующей вирусную ДНК, появляется короткий (4—6 п.н.) прямой повтор и, во-вторых, в местах «стыка» с клеточной ДНК оба вирусные LTR теряют по две концевые пары нуклеотидов.

Встраивание вирусного генома в клеточную хромосому — обязательная стадия репродукции ретровирусов независимо от того, обладают ли они онкогенным (трансформирующим) действием. Реплицируясь вместе с клеточной ДНК при митозе, вирус-специфическая ДНК — провирус — передается в дочерние клетки.

Синтез РНК ретровирусов происходит тогда, когда геном вируса в виде провирусной ДНК является интегральной частью клеточной хромосомы. Соответственно образование вирусных транскриптов идет в ядре и осуществляется клеточным транскрипционным аппаратом; в качестве основного фермента используется РНК-полимераза II. Поэтому большинство проблем, которые при этом нужно решить, — это обычные проблемы клеточной транскрипции (и посттранскрипционного процессинга), которые здесь описаны не будут. Но возникают и специфические проблемы.

Сравним структуру провируса и вирусной РНК (рис. 162). Провирус длиннее РНК, и транскрипция должна начинаться и оканчиваться внутри LTR: в одном LTR («левом») синтез РНК начинается, а в другом, структурно идентичном LTR («правом»), он терминируется. Это обстоятельство порождает несколько вопросов. Поскольку в «правом» LTR существует сигнал, обеспечивающий терминацию между последовательностями r и u_5 , почему такой же сигнал не «срабатывает» в «левом» LTR? Поскольку в «левом» LTR есть сигналы, обеспечивающие возможность инициации между u_3 и r , проявляются ли эти сигналы в «правом» LTR?

Промоторные элементы провируса расположены в районе u_3 ; таким образом, возможность транскрипции провируса возникает после появления района u_3 «впереди» вирусного ДНК-генома, т. е. после возникновения LTR. Примерно за 25 п. н. до стартовой точки транскрипции (до r) имеется характерный ТАТА-элемент, за 75 п. н. — СААТ-элемент и за 100—300 п. н. — энхансер. У разных ретровирусов энхансер имеет разную «силу», а у онкогенных ретровирусов «сила» энхансера может коррелировать со способностью вируса вызывать злокачественную трансформацию клеток-мишеней. Для активирования энхансера необходимо его взаимодействие с клеточными белками-регуляторами; в некоторых случаях, например у мышинового вируса рака молочных желез, эффективность энхансера регулируется гормонами (через посредство белков — рецепторов гормонов).

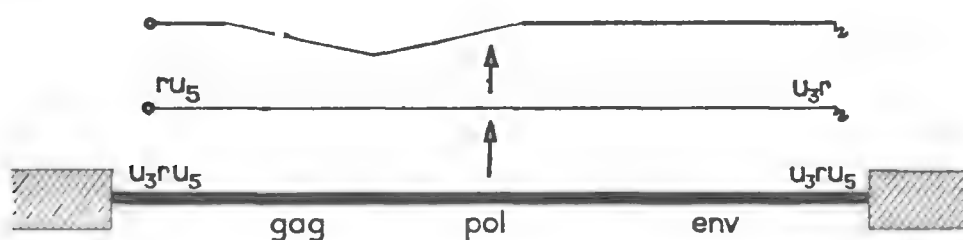


Рис. 162. Схема синтеза ретровирусных РНК

Внизу — провирус (жирная линия) в составе клеточной ДНК (заштриховано); на провирусе как на матрице синтезируется полноразмерная молекула вирус-специфической РНК (тонкая линия), которая затем может подвергаться сплайсингу и превращаться в субгеномную мРНК (показано образование только одного типа молекул мРНК); черный кружок — 5'-концевая кап-структура; волнистая линия — поли(А)

Так или иначе транскрипция провируса начинается с копирования «левого» района *r*; заканчивается же она после считывания «правого» *r*. По одной из гипотез, для терминации транскрипции провируса важно, чтобы в синтезируемой молекуле РНК возникла «шпилька» на границе районов *u3* и *r*; имеющаяся здесь нуклеотидная последовательность образование такой «шпильки» позволяет. Ясно, что при транскрипции «правого» *r* «шпилька» образоваться может, а при транскрипции «левого» *r* — не может, поскольку в образующейся здесь РНК район *u3* просто не представлен.

Нет полной ясности и в вопросе об инициации транскрипции в «правом» LTR, в котором, разумеется, имеются те же регуляторные сигналы, что и в «левом» LTR. Транскрипция в «левом» LTR начинается значительно чаще, чем в «правом». Есть предположение, что первый каким-то образом тормозит работу второго: по некоторым наблюдениям, инактивация «левого» LTR стимулирует инициацию транскрипции в «правом» LTR; в результате происходит считывание прилегающих клеточных генов с провирусного промотора.

У некоторых ретровирусов эта достаточно сложная картина еще более усложнена: в вирусном геноме могут быть дополнительно закодированы белки, регулирующие эффективность транскрипции провируса.

Первичные транскрипты провируса подвергаются обычным посттранскрипционным модификациям: экспонированию 5'-конца, полиаденилированию 3'-конца (в районе *r* есть сигнал полиаденилирования) и сплайсингу (рис. 162). Последняя модификация затрагивает не все транскрипции провируса — часть из них выходит в цитоплазму, сохраняя всю последовательность нуклеотидов. Такие молекулы РНК (помимо того, что они функционируют как мРНК для некоторых белков) включаются в вирион; тем самым завершается цикл репликации/транскрипции генома ретровирусов.

*Репликация/транскрипция геномов
вирусов гепатита В
и мозаики цветной капусты
(ДНК→РНК→ДНК)*

ДНК, входящая в состав частиц вируса гепатита В, — это молекула, построенная из двух линейных компонентов: полноразмерной (—)нити (~3,2 т. п. н.) с белком, ковалентно присоединенным к 5'-концу, а также сегмента (+)нити (1,7—2,8 т. п. н.). Этот сегмент содержит участки, комплементарные обоим концам (—)нити, и поэтому удерживает вирионную ДНК в кольцевой форме (рис. 163, а). В вирионе имеется вирус-специфическая ДНК-полимераза, способная достраивать (+)нить до размера полного генома. Геном вируса мозаики цветной капусты крупнее и содержит около 8 т. п. н.; это двухнитевая кольцевая молекула, обе цепи которой не непрерывны (рис. 163, б).

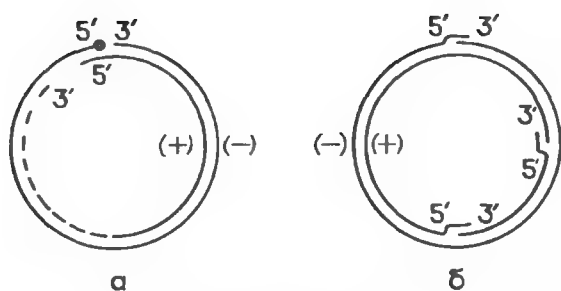


Рис. 163. Строение вирионной ДНК вирусов гепатита В (а) и мозаики цветной капусты (б)

Пунктирная линия — район, в котором (+)цепи вирионной ДНК могут иметь раз-
личную длину; черный кружок — белок,
присоединенный к 5'-концу (-) цепи

В зараженной клетке ДНК этих двух вирусов переходит в ковалентно-непрерывную форму, которая, как известно, удобна для репликации. Однако у обоих вирусов репликация ДНК-генома осуществляется при посредстве промежуточных линейных молекул РНК. Эти РНК образуются в результате транскрипции вирусных ДНК в клеточном ядре хозяйским ферментом РНК-полимеразой II. Транскрибируется только одна из нитей вирусной ДНК, причем

промоторы и терминаторы расположены на кольцевом геноме таким образом, что наряду с субгеномными мРНК образуются молекулы (+)РНК более длинные, чем геном. Ясно, что в длинных транскриптах должен быть прямой концевой повтор. Этот повтор способствует преодолению трудностей, возникающих при снятии ДНК-копии с 3'-конца линейной матрицы.

Дальнейшие этапы репликации в самом общем виде представляются следующим образом. Синтез ДНК на РНК-матрице происходит в результате обратной транскрипции, катализируемой вирус-специфической ДНК-полимеразой, которая способна использовать в качестве матрицы как ДНК, так и РНК, т. е. обладает свойствами обратной транскриптазы (ревертазы). Сначала синтезируется (-)нить ДНК; при этом в качестве затравки в случае вируса гепатита В используется белок (возможно, в виде нуклеотид-белкового комплекса), а в случае вируса мозаики цветной капусты — одна из клеточных тРНК. Затем на вновь синтезированной (-)нити ДНК тот же фермент строит (+)нить.

Таким образом, общая схема репликации / транскрипции генома этих вирусов поразительно похожа на схему репликации / транскрипции генома ретровирусов. По-видимому, это сходство имеет под собой эволюционную основу: во всяком случае, аминокислотные последовательности обратных транскриптаз всех трех групп вирусов (гепатита В, мозаики цветной капусты и ретровирусов) выявляют определенное сходство между собой.

4. РЕПЛИКАЦИЯ/ТРАНСКРИПЦИЯ ГЕНОМОВ РНК-СОДЕРЖАЩИХ ВИРУСОВ (РНК→РНК)

Виды вирусных РНК-геномов

Геномы почти всех известных РНК-содержащих вирусов — это линейные молекулы, которые в среднем короче вирусных ДНК-геномов. К самым крупным непрерывным РНК-геномам можно отнести РНК коронавирусов — ~28 т. н. Весьма крупными являются и геномы реовирусов — более 20 т. п. н. — однако здесь, как и во многих других случаях, генетическая информация распределена по нескольким (у реовирусов — по десяти) молекулам РНК.

Вирусные РНК-геномы удобно разделить на три группы. Во-первых, это *однонитевые геномы положительной полярности*, т. е. с нуклеотидной последовательностью, соответствующей таковой у мРНК (РНК фага Q_{β} , вирусов табачной мозаики, полиомиелита и др.). Во-вторых, это *однонитевые (—)РНК-геномы*; здесь нуклеотидные последовательности генома и мРНК взаимно комплементарны (например, у вирусов гриппа, бешенства, везикулярного стоматита и др.). Третью группу составляют *двухнитевые геномы* (реовирусы, вирусы цитоплазматического полиэдроза насекомых и др.). Способы репликации/транскрипции геномов этих трех групп вирусов существенно различны.

Известные двухнитевые РНК-геномы всегда сегментированы (т. е. состоят из нескольких разных молекул); в двух других группах вирусов имеются представители как с сегментированным, так и несегментированным геномом.

Обсуждая разнообразие форм вирусных РНК-геномов, целесообразно упомянуть о вириодах, которые вирусами в строгом смысле слова не являются. Вириоды — это низкомолекулярные (250—370 нуклеотидов) однонитевые кольцевые РНК, проявляющие свойства инфекционных патогенов растений.

Основные схемы репликации/транскрипции

а. Репликация / транскрипция однонитевых (+) РНК-геномов

Вирусные (+)РНК-геномы обычно кодируют несколько белков и часто вся генетическая информация содержится в единой молекуле. Если речь идет о геноме фага, то особых проблем с синтезом этих белков не возникает, так как в клетках прокариот *каждый цистрон* полицистронной матрицы *может транслироваться независимо*. Иначе обстоит дело у вирусов эукариот. В мРНК эукариот, как правило, функционирует только один иницирующий триплет. Чтобы

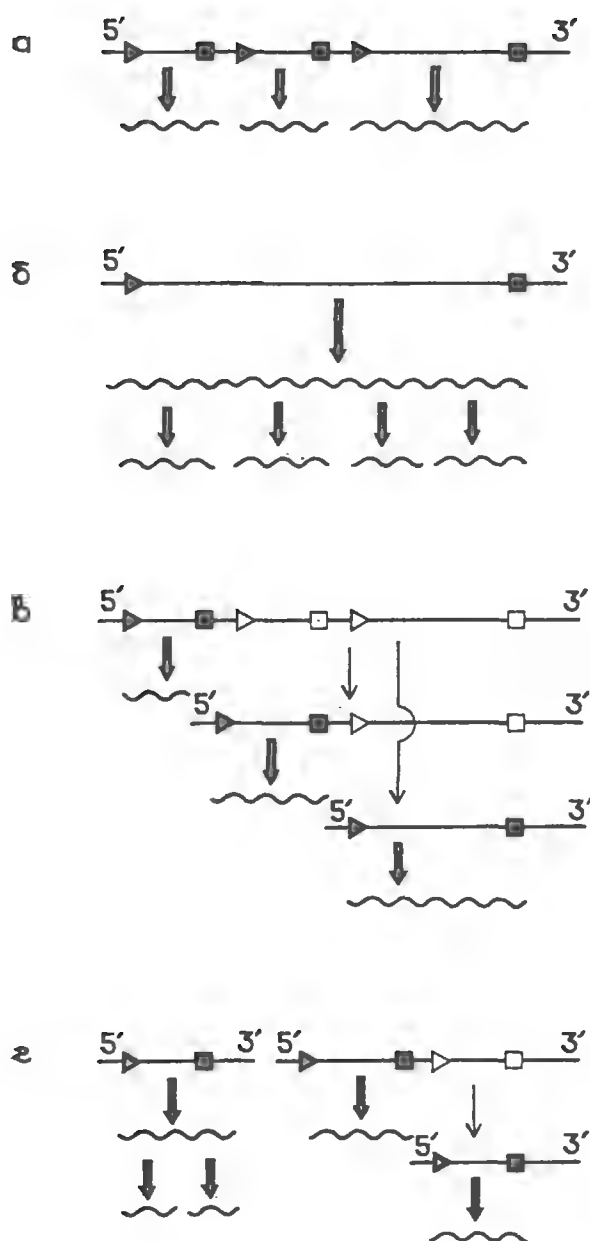


Рис. 164. Трансляционные «стратегии» вирусных (+) РНК-геномов:

а — независимая трансляция отдельных цистронов полицистронной (+)РНК; б — образование единого полипротеина и его последующее расщепление на функционально-активные белки; в — использование в качестве мРНК наряду с геномной (+)РНК также и субгеномных (+)РНК; г — при трансляции отдельных сегментов сегментированных (+)РНК-геномов может использоваться как механизм б, так и механизм в; прямые линии — (+)РНК; волнистые — полипептидные цепи: треугольники — инициаторные триплеты; квадраты — терминирующие триплеты (черные — реализуемые в данной молекуле РНК, светлые — потенциальные); тонкие стрелки — синтез РНК, промежуточные формы (—)РНК для простоты опущены; жирные — образование полипептидных цепей

образовать несколько белков, вирусам эукариот приходится полагаться на одну из следующих «стратегий» (рис. 164). Во-первых, в единой молекуле мРНК может быть закодирован *крупный полипептид-предшественник* — «полипротеин», из которого путем ограниченного протеолиза образуются «зрелые» белки. Во-вторых, цикл репликации/транскрипции вирусного генома может включать помимо полноразмерных (+) и (—)цепей также и более короткие — *субгеномные* — (+)РНК, локализация инициаторного триплета в которых отличается от локализации такого триплета в геномной (+)РНК. Наконец, геном может включать разные молекулы (+)РНК, т. е. быть *сегментированным*; однако эта последняя «стратегия» обычно не достигает своего логического завершения и количество (+)-сегментов остается меньше количества необходимых вирус-специфических белков. Поэтому последняя «стратегия» обычно комбинируется с образованием либо полипротеина-предшественника, либо субгеномных мРНК. Различия в трансляционных стратегиях неизбежно затрагивают и стратегии репликации / транскрипции вирусных (+) РНК-геномов.

Сначала рассмотрим системы, которые обходятся без образования субгеномных мРНК. Среди вирусов прокариот — это мелкие РНК-содержащие фаги, например фак

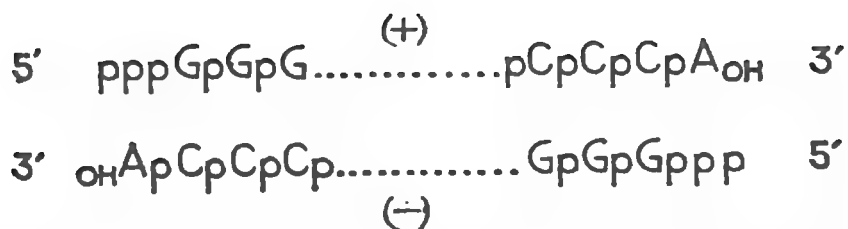


Рис. 165. Нуклеотидная последовательность концевых участков (+) и (—) цепей РНК фага Q_{β}

Q_{β} , а среди вирусов эукариот — пикорнавирусы, например вирус полиомиелита.

РНК фага Q_{β} содержит около 4 т. п. н. Попад в клетку, полицистронная фаговая РНК транслируется; в частности, образуется белок с $M_r \approx 65\,000$. Из клетки этот белок может быть изолирован в виде ферментативного комплекса — репликазы. Фермент содержит помимо вирус-специфического три клеточных белка. Удивительным образом эти три белка — компоненты трансляционного аппарата: белок S1 из малой субчастицы рибосом, а также два фактора элонгации полипептидной цепи — EF-Tu и EF-Ts. Комплекс этих четырех белков (одного вирус-специфического и трех клеточных) является РНК-зависимой РНК-полимеразой, но для эффективного использования в качестве матрицы (—)нити РНК фага Q_{β} необходим еще один клеточный полипептид — так называемый «хозяйский фактор»; этот низкомолекулярный белок ассоциирован с рибосомами, но его функция в незараженной клетке неизвестна (строго говоря, неизвестно, нужен ли «хозяйский фактор» для репликации фаговой РНК *in vivo*; даже *in vitro* потребность в нем исчезает после увеличения концентрации GTP в инкубационной среде).

Репликаза фага Q_{β} — высокоспецифичный фермент: из природных РНК он использует в качестве матрицы только собственный геном, т. е. РНК фага Q_{β} и некоторые родственные молекулы. Однако при наличии затравки фермент может копировать любую РНК: таким образом, специфичность проявляется на стадии инициации. Любопытно, что избирательность фермента по отношению к матрице в значительной степени определяется входящими в его состав клеточными белками: белок S1 и «хозяйский фактор» требуются при использовании в качестве матрицы (+)нити РНК фага Q_{β} , но эти белки не нужны, когда в роли матрицы выступает, например, (—)нить этой РНК.

Поскольку 3'-концевым нуклеозидом (—)нити РНК фага Q_{β} является аденозин, можно было бы ожидать, что первым (т. е. 5'-концевым) нуклеотидом в (—)нити будет UTP. На самом деле (—)нить начинается с GTP, который комплементарен предпоследнему остатку в (+)РНК фага Q_{β} (рис. 165). Точно такая же картина наблюдается и при синтезе (+)нити. Соответственно 3'-концевые аденозиновые остатки как (—)нити, так и (—)нити появляются в

этих молекулах нематричным путем. Эту работу выполняет сама репликаза Q_{β} .

Важная особенность репликационной системы фага Q_{β} заключается в том, что вновь синтезируемая нить не формирует протяженного дуплекса с матричной нитью; другими словами, на одонитевой матрице образуется одонитевый продукт. Предполагают, что возникновению межцепочечного дуплекса препятствует весьма выраженная внутрицепочечная вторичная структура матричной и вновь синтезируемой цепей.

Благодаря этой особенности, а также благодаря тому, что репликаза Q_{β} в присутствии «хозяйского фактора» распознает как (+)нить, так и (—)нить фаговой РНК, добавление этой РНК к ферменту запускает множественные повторные полные циклы репликации вирусного генома: на (+)матрицах синтезируются (—)нити, которые в свою очередь используются для синтеза (+)РНК. Соотношение между синтезом (+) и (—)цепей *in vitro* будет определяться концентрацией «хозяйского фактора», который, как известно, нужен для образования (—)цепей: при его недостатке предпочтительно идет синтез (+)нитей. Очищенная система, состоящая из репликазы Q_{β} , некоторого количества «хозяйского фактора», нуклеозидтрифосфатов и солей, способна синтезировать инфекционную фаговую РНК в количествах, многократно превышающих количество фаговой РНК, первоначально внесенное в систему в качестве матрицы.

Репликационная система вируса полиомиелита изучена менее детально; тем не менее здесь имеются явные отличия от только что рассмотренной фаговой системы. Так, на 5'-концах вновь синтезируемых (+) и (—)цепей полиовирусных РНК всегда присутствует низкомолекулярный вирус-специфический белок (VPg). Тирозиновый остаток VPg соединен фосфодиэфирной связью с 5'-концевым уридиловым остатком вирус-специфических РНК (обе комплементарные цепи начинаются с уридилового остатка). Большинство исследователей приписывают этому белку (или его комплексу с уридиловой кислотой) роль затравки при синтезе обеих нитей РНК; по этой точке зрения, VPg функционально аналогичен терминальному белку аденовирусов (см. раздел 1 этой главы).

Сам синтез осуществляется вирус-специфической РНК-полимеразой. Для работы очищенных препаратов этого фермента требуется не только матрица, но и затравка — комплементарный матрице олигонуклеотид; другими словами, фермент катализирует элонгацию цепи РНК, но не может осуществить инициацию этой цепи. Вновь синтезируемая цепь формирует межцепочечный дуплекс с матрицей. Из-за этого, а также из-за отсутствия соответствующих затравок очищенная РНК-полимераза вируса полиомиелита не способна осуществлять полный цикл репликации вирусного генома *in vitro*. Есть основания полагать, что и *in vivo* для репликации РНК вируса полиомиелита используется многокомпонентный аппарат, который включает помимо РНК-полимеразы и VPg по крайней мере еще

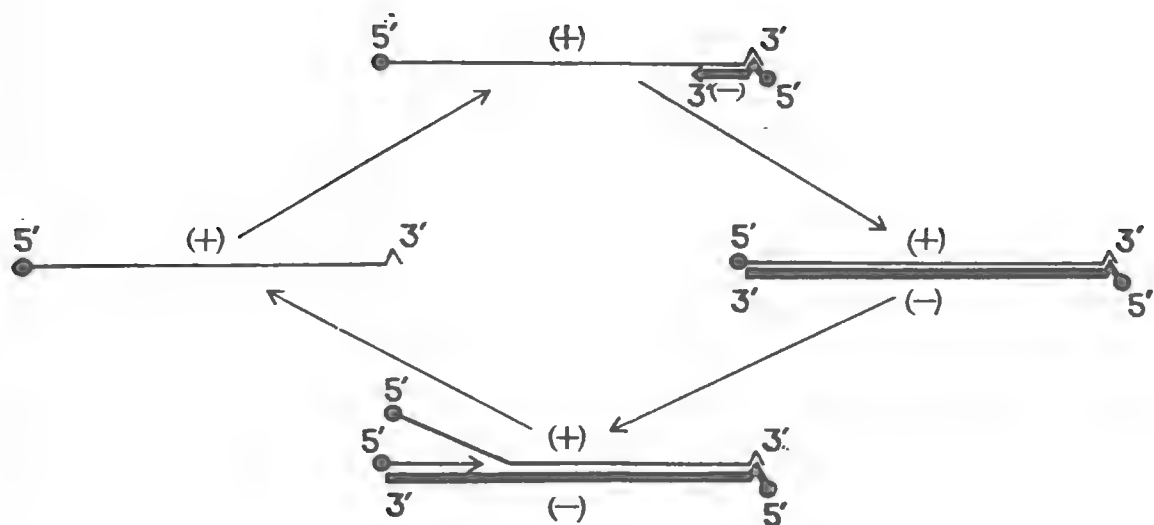


Рис. 166. Возможная схема репликации генома вируса полиомиелита
Кружок — терминальный белок. На 3'-конце (+)нитей и на 5'-конце (—)нитей находятся гомополимерные последовательности [поли(А) и поли(У), соответственно]

один вирус-специфический полипептид, а также и клеточные белки. Возможная схема репликации генома вируса полиомиелита приведена на рис. 166. Существенной чертой репликационного аппарата вируса полиомиелита является его связь с мембранами зараженной клетки; однако какую конкретно биологическую роль эта связь выполняет, пока не известно.

Наконец, необходимо указать, что РНК-геномы вируса полиомиелита и других пикорнавирусов способны участвовать в межмолекулярной рекомбинации. Скорее всего такая рекомбинация происходит по модели смены матриц: например, 3'-конец недостроенной (—)нити отделяется от матричной цепи и перемещается на гомологичное место другой оказавшейся по соседству молекулы вирусной РНК. После завершения синтеза (—)нити на новой матрице она может давать начало рекомбинантным (+)геномам.

Теперь о системах, в которых цикл репликации / транскрипции включает образование не только полноразмерных (+) и (—)нитей, но также и субгеномных мРНК. Эти субгеномные молекулы мРНК считываются с полноразмерной (—)нити; считывание может происходить двумя разными способами. В одних системах, например у альфа-вирусов или у вируса табачной мозаики (ВТМ), в (—)нити РНК имеются участки, которые можно уподобить промоторам: РНК-полимераза «узнает» эти участки и способна здесь (а не только на 3'-конце матрицы) начинать свою «работу». У альфа-вирусов образуется только один вид субгеномных мРНК (рис. 167), у других же (в том числе у ВТМ) — несколько. В последнем случае субгеномные мРНК разной длины имеют разные 5'-концы, поскольку их синтез начинается в разных местах (—)матрицы. В то же время 3'-концы всех субгеномных мРНК, образованных на данной матрице, идентичны и соответствуют 3'-концу полноразмерной (+)нити.

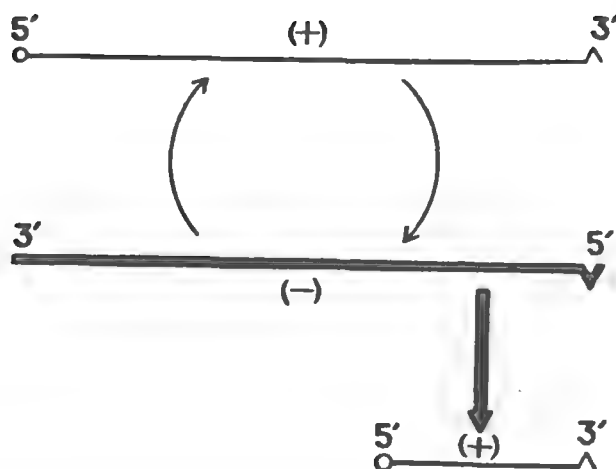


Рис. 167. Схема репликации/транскрипции генома альфавирусов

На 5'-конце геномной и субгеномной (+)РНК имеется *кэл* (кружок); на 3'-конце молекул (+)РНК и на 5'-конце молекул (-)РНК находится гомополимерные последовательности [поли(А) и поли(У) соответственно]; жирная стрелка — транскрипция, тонкие — репликация

Другой вариант образования субгеномных мРНК реализуется у коронавирусов. В зараженной клетке помимо геномной РНК синтезируется шесть классов субгеномных мРНК. Нуклеотидные последовательности всех этих (+)РНК идентичны не только на 3'-конце, но и на 5'-конце. Объясняется это следующим образом (рис. 168). Как и в других системах, сначала с геномной РНК считывается (-)нить. Синтез всех (+)цепей начинается на 3'-конце (-)матрицы. Однако связь между синтезируемой (+)нитью и матричной (-)ни-

тью весьма слаба. После образования 5'-концевого «лидерного» сегмента (+)нити длиной около 70 нуклеотидов этот «лидер» имеет высокий шанс отделиться от матрицы. В дальнейшем он опять взаимодействует с той же или другой молекулой (-)РНК. Однако, поскольку в матрице имеются характерные прямые повторы, 3'-конец лидерного сегмента может присоединиться либо к тому месту, от-

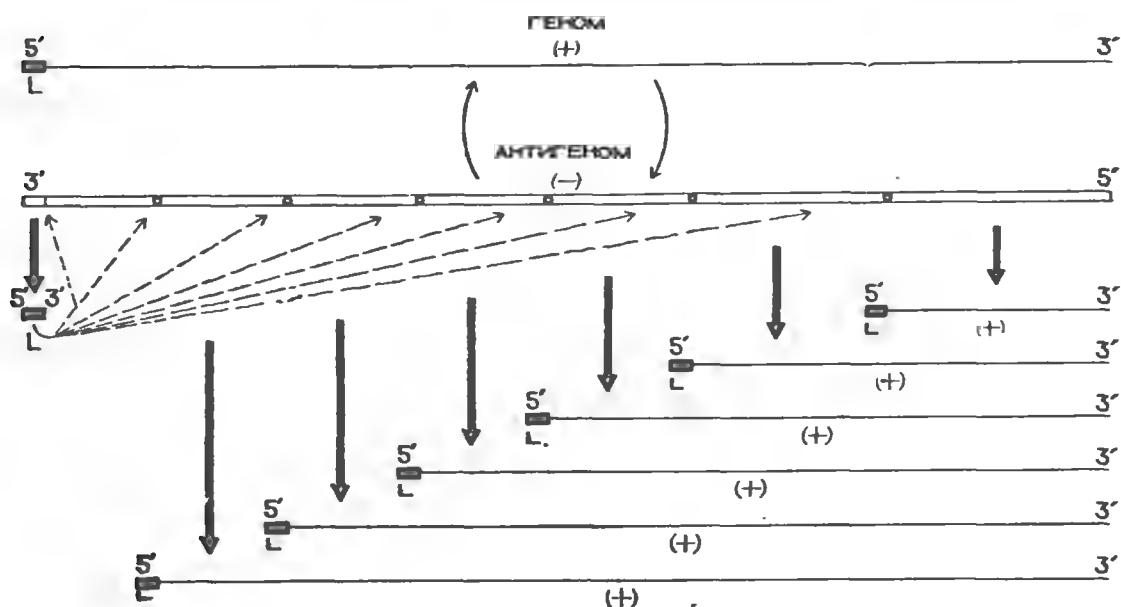


Рис. 168. Схема репликации/транскрипции генома коронавирусов:

Жирный прямоугольник — 5'-концевая лидерная (L) нуклеотидная последовательность (+)нитей; кружки в (-)нити (антигеноме) — участки, комплементарные 3'-концу лидерного сегмента; жирные стрелки — транскрипция; тонкие — репликация. Отсоединившийся от (-)матрицы лидерный сегмент может снова присоединиться к одному из комплементарных участков этой матрицы (показано прерывистыми стрелками) и послужить в качестве затравки для синтеза разных субгеномных мРНК

куда он «ушел», либо к другому гомологичному месту. Далее синтез (+)нити возобновляется, и в зависимости от выбора «места посадки» лидера возникает либо геномная, либо одна из субгеномных (+)РНК. Такая легкость отсоединения вновь синтезируемой цепи от матрицы также обуславливает высокий уровень рекомбинации между РНК-геномами коронавирусов; эта рекомбинация осуществляется, очевидно, по модели смены матрицы.

Естественно, что помимо субгеномных (+)РНК одним из продуктов репликации / транскрипции должны быть и полноразмерные (геномные) (+)РНК, которые, во-первых, направляют синтез белков, закодированных в 5'-концевом районе генома, а во-вторых, включаются в дочерние вирусные частицы. Полноразмерные (+)РНК считываются с такой же (—)матрицы, как и субгеномные мРНК. Динамика образования различных видов вирус-специфических РНК различна: синтез (—)РНК более характерен для ранних стадий инфекционного цикла, а синтез (+)РНК — для поздних; обнаруживаются и различия в динамике синтеза субгеномных и полноразмерных (+)РНК. Известно, что в этой регуляции принимают участие вирус-специфические белки, но конкретные их функции пока не выяснены, если не считать, что некоторые из них входят в состав РНК-зависимой РНК-полимеразы.

5'-концы геномных и субгеномных (+)РНК вирусов эукариот часто кэпированы. Соответствующая посттранскрипционная модификация вирусных РНК осуществляется клеточными ферментами.

б. Репликация / транскрипция однонитевых (—)РНК-геномов

Многие вирусы имеют геном в виде (—)нити РНК. У некоторых таких вирусов геном представлен единой непрерывной молекулой, а у других он сегментирован, т. е. состоит из нескольких молекул. Общим свойством вирусов с (—)РНК-геномом является то, что в состав их вирусных частиц входит РНК-полимераза, способная копировать РНК-матрицу. Биологический смысл такой организации понятен. Поскольку, по определению, (—)РНК не может выполнять функции мРНК, для образования «своих» мРНК вирус должен внести в клетку не только геном, но и фермент, умеющий снимать с этого генома комплементарные копии. Другое общее свойство этих вирусов заключается в том, что матрицей для репликации / транскрипции является не свободная РНК, а вирусный рибонуклеопротеид (РНП) — молекула РНК, равномерно покрытая вирус-специфическим белком.

Сначала — о вирусах с непрерывным (—)РНК-геномом. Наиболее изученный представитель этой группы — вирус везикулярного стоматита (ВВС). Геном ВВС — однонитевая РНК, содержащая около 11 т. н. В вирусной частице эта РНК формирует РНП, комплексируясь с белком N. С вирусным РНП прочно ассоциированы

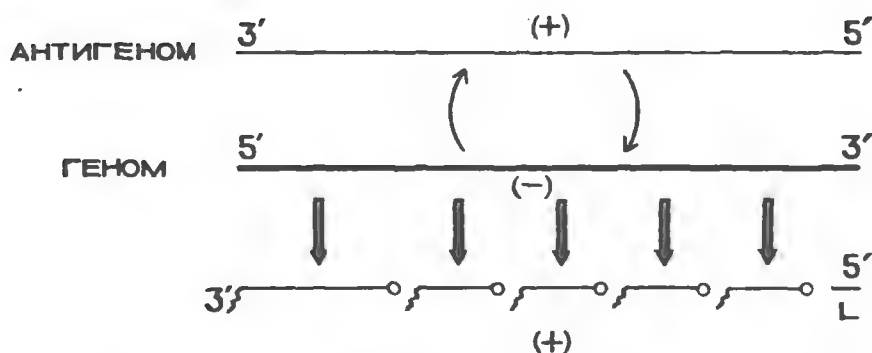


Рис. 169. Схема репликации/транскрипции генома вируса везикулярного стоматита:

L — лидерный сегмент (+)РНК: кружки — 5'-концевые кэп-структуры; волнистые линии — 3'-концевые поли(А)-последовательности; жирные стрелки — транскрипция; тонкие — репликация

также молекулы РНК-полимеразы и еще одного белка, принимающего участие в синтезе РНК. После взаимодействия с чувствительной клеткой субвирусная частица, содержащая все эти белки и геномную РНК, освобождается от наружной липопротеидной оболочки и оказывается в цитоплазме. Сразу же начинается синтез молекул мРНК. Эти мРНК моноцистронны — им ведь предстоит функционировать в клетках эукариот. Хотя многие важные детали механизма образования моноцистронных (+)РНК на непрерывной (—)РНК ВВС еще не установлены, наиболее вероятная модель этого процесса такова (рис. 169).

РНК-полимераза, ассоциированная с вирусным РНП, начинает транскрипцию, по-видимому, непосредственно на 3'-конце геномной (—)РНК и передвигается к ее 5'-концу. Однако по ходу этого движения фермент встречает специальные сигналы, обуславливающие терминацию и реинициацию цепи (+)РНК. Первый такой терминирующий сигнал расположен недалеко от 3'-конца (—)РНК-матрицы. В результате образуется короткая (около 50 нуклеотидов) так называемая «лидерная» (+)РНК, которая не транслируется. Более интересные и сложные события происходят на границах между участками (—)РНК, кодирующими индивидуальные белки (рис. 170). Каждый такой участок заканчивается (если идти от 3'- к 5'-концу матрицы) последовательностью из семи остатков уридиловой кислоты. Соответственно в синтезируемой (+)цепи в этом месте появляются остатки адениловой кислоты, но не семь таких остатков, как можно было бы ожидать, а несколько десятков или даже сотни. РНК-полимераза здесь как бы «буксует», образуя 3'-концевой поли(А) сегмент моноцистронной мРНК. После завершения образования поли(А) фермент возобновляет копирование (—)РНК, но реинициация следующей молекулы мРНК происходит на третьем от олигоуридилового тракта нуклеотиде матрицы. В результате два «межгенных» нуклеотида геномной РНК оказываются нетранскрибированными. 5'-концы молекул мРНК кэпируются; это происходит,

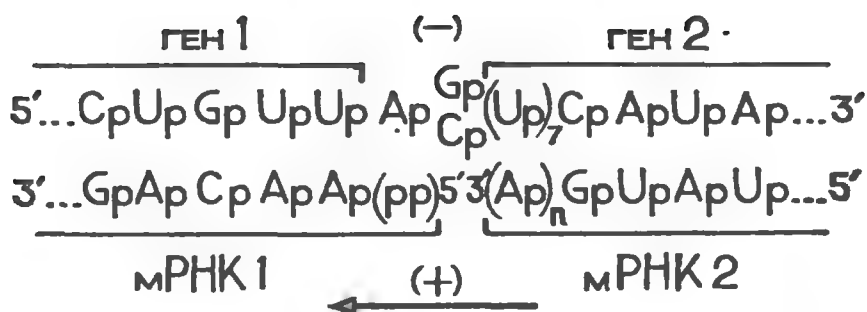


Рис. 170. Схема событий, происходящих на границе между двумя генами при транскрипции генома вируса везикулярного стоматита

Сверху — последовательность вирусной (—)РНК в области, где соседствуют два гена (в состав межгенового промежутка наряду с адениловым остатком могут входить либо гуаниловый, либо цитидиловый остатки); снизу — структура концов двух мРНК; стрелкой указано направление транскрипции

по-видимому, при участии вирус-специфических белков. Все эти события повторяются на границе между каждой парой генов.

После трансляции вновь синтезированных мРНК и накопления соответствующих белков начинается собственно репликация генома ВВС. Сначала синтезируются точные, полноразмерные (+)копии вирусного генома. Для этого необходимо подавить «буксование» РНК-полимеразы на полиуридилиловых последовательностях матрицы, а также внутреннюю терминацию. Предполагают, что такое регуляторное переключение происходит в результате взаимодействия вирус-специфических белков (вероятно, белка N) с растущей (+)цепью. Во всяком случае, все имеющиеся в зараженной клетке полноразмерные молекулы (+)РНК находятся там в виде РНП, сходного по структуре с РНП, содержащим геномную (—)РНК. В заключение на полноразмерной (+)РНК синтезируются (—)нити, которые включаются в состав дочерних вирионов.

Репликацию / транскрипцию сегментированных (—)РНК-геномов удобно рассмотреть на примере вируса гриппа (рис. 171). Этот вирус содержит восемь разных молекул (—)РНК длиной от 900 до

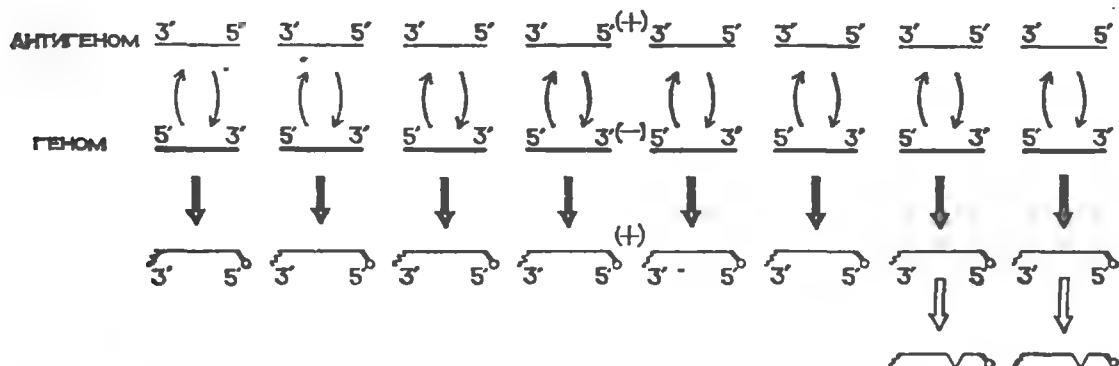


Рис. 171. Схема репликации/транскрипции генома вируса гриппа:

Горизонтальные линии — вирус-специфические нуклеотидные последовательности; кружки и примыкающие с ним косые отрезки — 5'-концевые кэп-структуры и олионуклеотиды, происходящие из клеточных транскриптов; волнистые линии — 3'-концевые поли(А)-последовательности; жирные стрелки — транскрипция; тонкие стрелки — репликация; светлые стрелки — сплайсинг

2350 нуклеотидов (всего 13,5 т. н.). 5'- и 3'-концы каждого сегмента частично взаимно комплементарны; при этом одноименные концы всех сегментов имеют одинаковые нуклеотидные последовательности. Каждый из сегментов РНК покрыт главным структурным белком; с этим РНП ассоциированы три относительно крупных полипептида — так называемые Р белки, которые осуществляют синтез вирус-специфических РНК.

Вскоре после инфекции вирусный нуклеопротеид оказывается в ядре зараженной клетки, где начинается образование вирусных мРНК. Отметим, что ядерная локализация синтеза вирусных РНК необычна для РНК-содержащих вирусов. Необычен и механизм инициации молекул мРНК. В качестве затравки здесь используются фрагменты транскриптов клеточных генов. Один из вирусных Р-белков «узнает» экпированные 5'-концы этих транскриптов и фиксирует их в районе 3'-концов геномных РНК. Затем этот же или другой Р-белок вносит эндонуклеолитический разрыв в клеточный транскрипт на расстоянии 10—13 нуклеотидов от его 5'-конца. Возникающий при этом короткий экпированный клеточный олигонуклеотид и используется как затравка при синтезе (+)нити. Таким образом, на 5'-концах мРНК вируса гриппа появляется гетерогенный набор олигонуклеотидов клеточного происхождения. При образовании вирусных мРНК (—)РНК-матрица копируется не полностью. На расстоянии около 20 нуклеотидов от 5'-конца в ней находится олигоуридиловая последовательность длиной 5—7 нуклеотидов. Как и в случае ВВС, на этом месте РНК-полимераза вируса гриппа «буксует» и образует 3'-концевой полиадениловый тракт молекул мРНК. После завершения синтеза поли(А) происходит терминация и около двух десятков 5'-концевых нуклеотидов геномных РНК не представлены в вирус-специфических мРНК. Из образовавшихся таким образом восьми видов вирусных мРНК шесть кодируют по одному белку. Однако часть молекул двух других мРНК подвергается сплайсингу, осуществляемому клеточными ферментами по обычным механизмам. В возникающих в результате сплайсинга дополнительных видах мРНК используются альтернативные открытые рамки считывания, подобно тому, как это характерно для ДНК-содержащих вирусов.

Таким образом, при использовании восьми сегментов геномной (—)РНК в зараженной вирусом гриппа клетке образуется 10 классов моноцистронных мРНК. Эти мРНК, переходя в цитоплазму, направляют синтез вирус-специфических белков, но по очевидным причинам они не могут быть использованы в качестве матриц для обратного синтеза геномных (—)РНК: молекулы мРНК на 5'-конце содержат гетерогенные фрагменты транскриптов клеточных генов, а их 3'-концы полиаденилированы и лишены комплемента 5'-конца геномных (—)РНК. Для репликации генома необходимы точные и полные (+)РНК-копии всех сегментов вирусного генома. Такие некэпированные и неполиаденилированные копии действительно

появляются в зараженной клетке после накопления достаточного количества вирус-специфических белков. Механизм, обеспечивающий переключение с синтеза мРНК на образование полноразмерных (+)РНК, пока не расшифрован, но, как и в случае ВВС, он должен включать угнетение «буксования» РНК-полимеразы на олигоуридиловых участках матрицы и подавление преждевременной термినации; кроме того, должен измениться способ инициации синтеза (+)нити (клеточные РНК здесь не нужны). Полноразмерные (+)РНК входят в состав вирус-специфических РНП и используются в качестве матриц для завершения цикла репликации — для образования геномных (—)РНК.

Известны и другие вирусы как с непрерывным, так и сегментированным (—)РНК-геномом. И хотя какие-то детали механизмов их репликации / транскрипции могут отличаться от только что описанных, общность основных принципов прослеживается удивительно хорошо. Но на одной группе вирусов следует остановиться особо. Речь идет о флебовирусах, которые относят к обширному семейству буньявирусов. Все буньявирусы имеют геном, состоящий из трех сегментов. У «обычных» буньявирусов, подобно вирусу гриппа, все эти сегменты представляют собой молекулы (—)РНК, на каждой из которых как на матрице синтезируются как субгеномные мРНК с гетерогенными экспонированными 5'-концами, происходящими из клеточных РНК, так и полноразмерные молекулы (+)РНК. Здесь рассмотрим в первую очередь малый (S) сегмент буньявирусов; у обычных представителей этого семейства он содержит информацию о двух белках, которые считываются с одной молекулы субгеномной мРНК благодаря использованию разных инициаторных триплетов и разных открытых рамок считывания. Замечательная особенность флебовирусов, которые во всех остальных отношениях являются типичными буньявирусами, заключается в том, что S-сегменту их генома нельзя приписать определенный знак: открытая рамка считывания для одного из белков находится в геномной РНК, а для другого — в комплементарной ей РНК.

Таким образом, 5'-половина S-сегмента генома флебовирусов соответствует (+)полярности, а 3'-половина — (—)полярности. Такие РНК можно назвать *обоюдозначащими*. Схема репликации / транскрипции S-сегмента флебовирусов выглядит следующим образом (рис. 172). С геномной РНК считываются два класса комплементарных РНК: во-первых, субгеномная мРНК [(+)РНК], кодирующая один из белков, и, во-вторых, полноразмерная комплементарная обоюдозначащая РНК — антигеном. В свою очередь, эта комплементарная РНК (антигеном) служит матрицей для образования также двух видов РНК: второй субгеномной мРНК [(+)РНК], кодирующей другой белок, и полноразмерной обоюдозначащей геномной РНК, которая идет на построение нового поколения вирусных частиц. Что касается других сегментов генома флебовирусов, то это обычные (—)РНК.

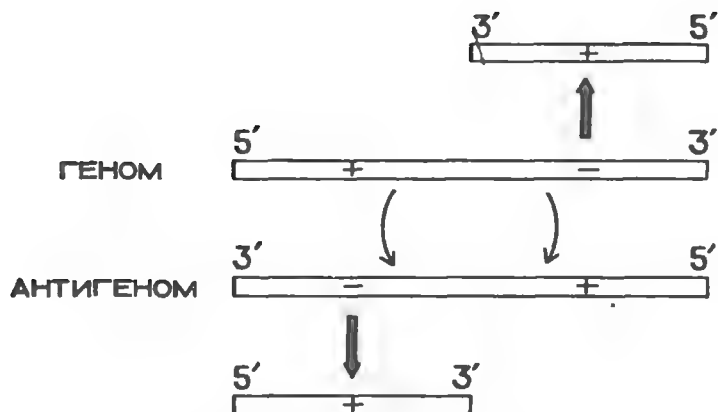


Рис. 172. Схема репликации/транскрипции S-сегмента генома флебовирусов:

жирные стрелки — транскрипция; тонкие — репликация

Как ни необычна тактика репликации / транскрипции S-сегмента генома флебовирусов, она не уникальна. Сходная картина наблюдается и в случае одного из двух сегментов генома другой группы вирусов — аренавирусов.

в. Репликация / транскрипция двухнитевых РНК-геномов

Двухнитевые РНК-геномы встречаются как у вирусов эукариот, так и у вирусов прокариот. Системы репликации / транскрипции у разных представителей этой группы вирусов могут заметно различаться. Рассмотрим, как эти процессы осуществляются у реовирусов (рис. 173).

После заражения клетки наружная белковая оболочка вируса повреждается и в цитоплазме оказывается модифицированная вирусная частица. Имеющиеся в клетке рибонуклеозидтрифосфаты получают доступ к вирусной сердцевине, которая содержит 10 сегментов двухнитевой геномной РНК и несколько белков, в том числе РНК-полимеразу, способную использовать РНК-дуплекс в качестве матрицы. Начинается асимметрический и консервативный синтез

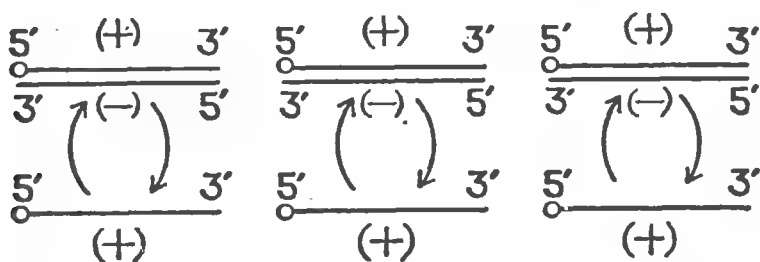


Рис. 173. Схема репликации/транскрипции генома реовируса:

кружки — 5'-концевые кэп-структуры. На рисунке изображены 3 из 10 сегментов генома реовируса

РНК. Это означает, что образуются только (+)нити, и при этом предшествующие родительские (+)нити не вытесняются из дуплекса. 5'-концы вновь синтезируемых молекул РНК кэпируются, по-видимому, при участии вирусных ферментов. Полиаденилирование 3'-концов (+)нитей не происходит.

На первых порах (+)нити выполняют роль мРНК: они направляют образование вирус-специфических белков. После накопления достаточного количества этих белков начинается формирование субвирусных частиц. При этом в одну субвирусную частицу, содержащую некоторые из вирус-специфических белков, включается полный набор, т. е. 10 разных видов, молекул (+)РНК. Механизм такого избирательного и организованного белок-нуклеинового взаимодействия пока не понятен. Вирус-специфическая РНК-полимераза является интегральным компонентом субвирусных частиц и осуществляет синтез двухнитевых РНК-геномов, используя в качестве матрицы находящиеся в этих же частицах (+)нити РНК. После того как РНК субвирусной частицы переходит в двухнитевую форму, может опять начаться синтез однонитевых (+)РНК. Но если к этому времени в клетке накопилось достаточно много белков, необходимых для построения наружной оболочки вируса, то формируются зрелые вирионы, в которых дальнейший синтез РНК блокируется.

Таким образом, в отличие от некоторых ранее рассмотренных вирусных систем у реовирусов в качестве матриц как для синтеза белка, так и для синтеза (—)РНК используются одинаковые молекулы (+)РНК. Кроме того, у реовирусов не наблюдается каких-то особых проблем с инициацией цепей РНК. Отметим, впрочем, что у некоторых других вирусов с двухнитевым РНК геномом (например, у бирнавирусов, поражающих насекомых) в инициации цепей РНК, возможно, участвует белок, который оказывается ковалентно-связанным с 5'-концом геномной РНК. Отметим также, что у некоторых других вирусов этой группы (в частности, у фага ф6) синтез (+)нитей на родительском РНК-дуплексе происходит по полуконсервативной модели; синтез новой (+)нити всегда сопряжен с вытеснением из дуплекса предсуществующей (+)нити.

г. Репликация РНК вирионов и вирусоидов

Вироиды — это инфекционные низкомолекулярные (от 250 до 370 нуклеотидов) однонитевые кольцевые РНК, способные к автономному размножению в растениях и обычно вызывающие развитие патологических симптомов, например веретеновидность клубней у картофеля, тяжелое заболевание каданг-каданг кокосовых пальм и др. Наиболее важное отличие вирионов от вирусных геномов — это то, что первые не кодируют никаких специфических белков. Тем не менее для удобства изложения будем обозначать вирионную РНК

знаком «+» (здесь этот знак указывает лишь на то, что мы имеем дело с основной инфекционной формой вирионной РНК).

Репликация вирионной РНК происходит в ядре зараженной клетки; вероятная схема этого процесса такова (рис. 174). Сначала на кольцевой (+)матрице синтезируется комплементарная (—)цепь. Этот синтез осуществляется клеточным ферментом; в качестве одного из кандидатов рассматривают ДНК-зависимую РНК-полимеразу II. Возможно, расширению специфичности этого фермента, обычно использующего двухнитевую ДНК-матрицу, способствует то обстоятельство, что вирионная РНК содержит необычно высокую (для одонитевых нуклеиновых кислот) долю элементов с вторичной структурой. Синтез идет, вероятно, по модели разматывающегося рулона (см. раздел 1 этой главы), и в результате появляются линейные олигомерные (—)нити. Затем происходит образование линейных олигомерных (+)нитей; не ясно, используются ли при этом в качестве матрицы олигомеры (—)нитей или образовавшиеся из них кольцевые молекулы. Далее линейные (+)олигомеры превращаются в кольцевые мономерные молекулы — конечный продукт реплика-

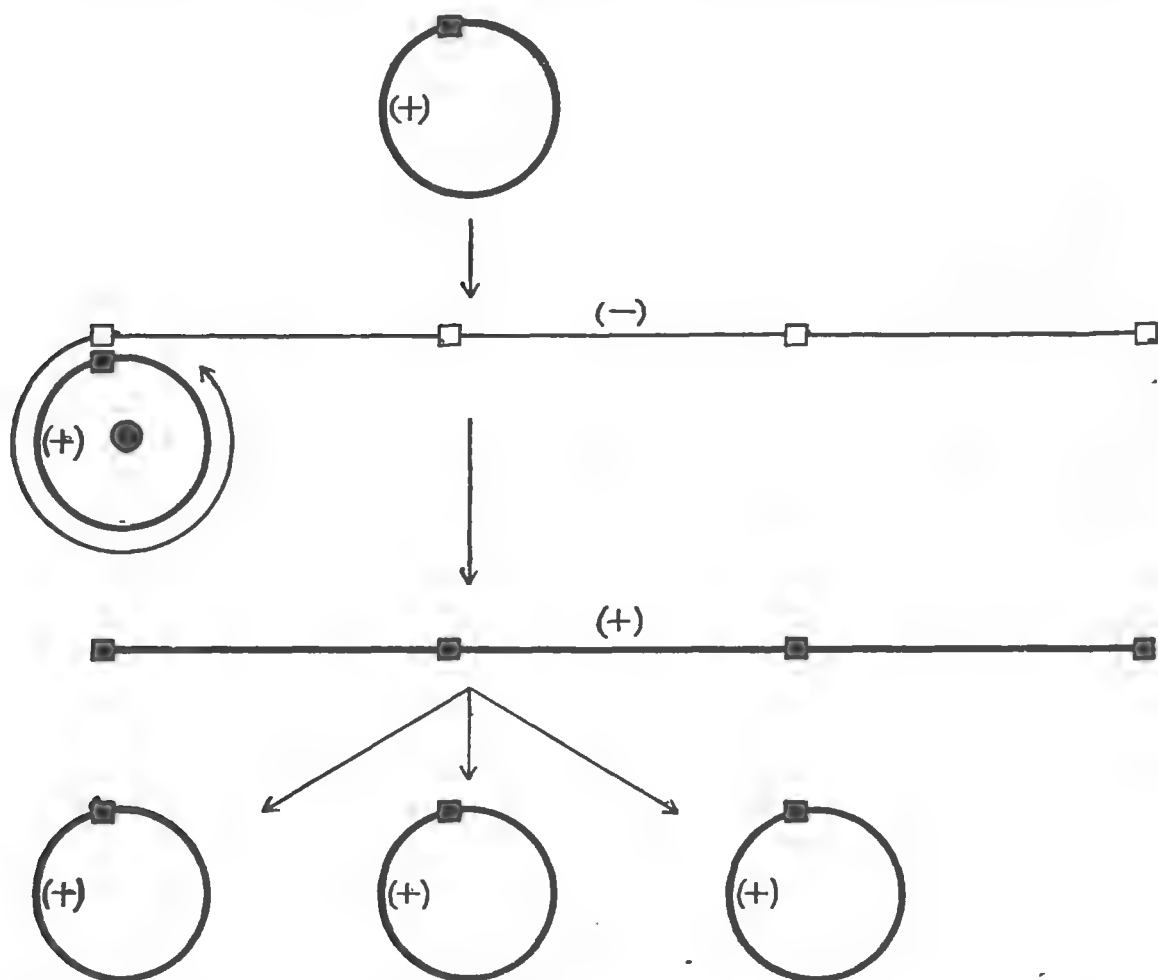


Рис. 174. Схема репликации вирионов

Жирные линии — (+)нити; тонкие — (—)нити; квадраты — участки, в которых происходит разрыв конкатемера, сопряженный с циркуляризацией мономерных РНК

ции. В ряде систем превращение «линейный олигомер $\longrightarrow$ кольцо» осуществляется по механизму «самосплайсинга» (см. гл. VIII), т. е. автокаталитического выщепления и одновременной циркуляризации участков РНК, расположенных между определенными структурированными элементами молекулы. Теоретически олигомер может также сначала расщепляться на линейные мономеры под действием клеточных эндонуклеаз (или автокаталитически), а затем эти мономеры будут переходить в кольцевую форму при помощи имеющихся в клетке РНК-лигаз.

«Урожай» вирионной РНК в клетке может достигать 10^5 молекул.

Здесь же следует упомянуть и о вирусоидах — однопитевых кольцевых РНК сходных с вирионами размеров, обнаруживаемых в препаратах некоторых вирусов растений. В отличие от вирионов вирусоидная РНК не способна к автономной репликации в клетке — ее синтез происходит только в присутствии «обычного» вируса, который, по-видимому, предоставляет вирусоиду свою РНК-полимеразу. В остальном схемы репликации вирионных и вирусоидных РНК очень похожи, если не идентичны. Вирусоид не может размножаться без вируса-помощника, но вирус без вирусоида обходится вполне благополучно; правда, при этом могут измениться вызываемые вирусом патологические симптомы.

* * *

Были рассмотрены различные способы репликации / транскрипции вирусных РНК-геномов. Кажется целесообразным сказать хотя бы несколько слов о биологическом смысле этого разнообразия.

Наиболее простой цикл репликации / транскрипции вирусной РНК — это когда с геномной РНК снимается комплементарная копия и эта копия, в свою очередь, служит матрицей для синтеза геномной РНК; роль мРНК в образовании всех необходимых для размножения вируса белков выполняет родительская РНК. Если отвлечься от частных случаев, то этот принцип реализуется у фага Q_{β} и у вируса полиомиелита. Однако стратегии этих вирусов различаются в одном существенном отношении. Фаг Q_{β} размножается в клетках прокариот, поэтому его (+)РНК может функционировать как истинная полицистронная мРНК. Хозяин вируса полиомиелита — эукариотная клетка. Соответственно на (+)РНК этого вируса имеется единственная точка инициации трансляции, и все «зрелые» вирус-специфические белки возникают в результате ограниченного протеолиза единого полипротеина-предшественника. Как и у ДНК-содержащих вирусов, у вирусов с РНК-геномом разные вирус-специфические белки требуются в разных количествах и в разное время, а образование всех этих белков из единого предшественника затрудняет количественную и временную регуляцию их производства. Поэтому у РНК-содержащих вирусов эукариот возникли механизмы, обеспечивающие появление разных мРНК для

разных белков — образование субгеномных мРНК и сегментация генома у вирусов с (+)РНК-геномом.

Вирусы с (+)РНК-геномами относительно просто устроены, и это их большое достоинство. Однако при естественных условиях заражения в клетку часто попадает лишь одна молекула вирусной (+)РНК. Эта единственная молекула должна выдержать конкуренцию за компоненты белок-синтезирующего аппарата с «преобладающими силами» клеточных мРНК да еще и сохраниться, чтобы послужить матрицей для репликации вирусного генома. Поэтому кажется естественным возникновение вирусов, у которых геном представляет собой не (+)РНК, а матрицу для ее образования, т. е. (—)РНК или двухнитевую РНК. Такие вирусы должны быть более сложно устроены; в частности, они должны уметь вводить в клетку не только вирусный геном, но и заранее заготовленный фермент, нужный для синтеза молекул мРНК. Но зато такие системы приобретают очень большое преимущество — попавшая в клетку единственная вирусная частица способна наработать огромное количество молекул мРНК. Если речь идет о вирусах эукариот, то опять кажется разумным обеспечить образование нескольких видов мРНК, даже если родительский геном — это единая молекула.

Таким образом, каждая стратегия имеет свои достоинства и недостатки и сам факт существования разнообразных РНК-содержащих вирусов свидетельствует о том, что эти достоинства и недостатки хорошо сбалансированы.

ЛИТЕРАТУРА

Агол В. И. Разнообразие и единство способов репликации вирусных ДНК // Усп. соврем. биол. 1986. Т. 20. С. 309—328.

Грен Э. Я. Регуляторные механизмы репликации РНК-содержащих бактериофагов.— Рига: Зинатне, 1974.

Лурия С., Дарнелл Дж., Балтимор Д., Кэмпбелл Э. Общая вирусология.— М.: Мир, 1981.

Пташник М. Переключение генов.— М.: Мир, 1989.

Хесин Р. Б. Непостоянство генома.— М.: Наука, 1984.

Baas P. D. DNA replication of single-stranded Escherichia coli DNA phages // Biochim. biophys. acta. 1985. V. 825. P. 111—139.

Bacteriophage T4/Eds. Mathews C. K., Kutter E. M. Mosig G., Berget P. B./ Am. Soc. Microbiol., Wash., 1983.

Banerjee A. K. Transcription and replication of rhabdoviruses // Microbiol. Rev. 1987. V. 51. P. 66—87.

Berns K. I., Bohenzky R. A. Adeno-associated viruses: An update // Adv. Virus Res. 1987. V. 32. P. 243—306.

Bishop D. H. L. Ambisense RNA genomes of arenaviruses and phleboviruses // Adv. Virus Res. 1986. V. 31. P. 1—51.

Cotmore S. F., Tattersall P. The autonomously replicating parvoviruses of vertebrates // Adv. Virus Res., 1987. V. 33. P. 91—174.

DNA Tumor Viruses/Eds. Botchan M., Grodziker T., Sharp P. N.-Y., Cold Spring Harbor Laboratory, 1986.

Flint S. J. Regulation of adenovirus mRNA formation // Adv. Virus Res., 1986. V. 31. P. 169-228.

Fundamental Virology/Ed. Fields. B. N.-Y.: Raven Press, 1986.

Ganem D., Varmus H. E. The molecular biology of the hepatitis B virus // *Ann. Rev. Biochem.*, 1987. V. 56. P. 651—693.

Hirth L. The molecular biology of caulimoviruses // *Microbiol. Sci.*, 1986. V. 3. P. 260—265.

Joklik W. K. Recent progress in reovirus research // *Ann. Rev. Genet.*, 1985. V. 19. P. 536—575.

Lai M. M. C. Coronavirus leader-RNA-primed transcription: an alternative mechanism to RNA splicing // *BioEssays*, 1986. V. 5. P. 257—260.

Lambda II/Eds. Hendrix R. W., Roberts J. W., Stahl F. W., Weisberg B. A. N.-Y., Cold Spring Harbor Laboratory, 1983.

Mason W. S., Taylor J. M., Hull R. Retroid virus genome replication // *Adv. Virus Res.*, 1987. V. 32. P. 35—96.

Mizuuchi K., Craigie R. Mechanism of bacteriophage Mu transposition // *Ann. Rev. Genet.*, 1986. V. 20. P. 385—429.

Riesner D., Gross H. J. Viroids // *Ann. Rev. Biochem.*, 1985. V. 54. P. 531.

The Molecular Biology of Viral Replication/Ed. Perez-Bercoff R. N.-Y.: Plenum. 1987.

Vartapetian A. B., Bogdanov A. A. Proteins covalently linked to viral genomes // *Prog. Nucl. Acid Res. a. Mol. Biol.*, 1987. V. 34. P. 209—251.

- Авторадиография 17, 215
 Агнопротеин 299, 302, 303
 Аденилирование 171
 Адениловая РНКаза 39
 Аденин 5, 6, 11, 12, 21, 38, 73, 82
 129, 131, 143
 Аденовирусы 261, 264, 267—268, 320
 — транскрипция генома 303—306
 S-аденозилметионин 130—132
 Аденозин 12, 23, 177, 319
 Аденозинмонофосфат 55
 Аденозинтрифосфат 49—54, 57, 82,
 132, 179. См. также Гидролиз
 аденозинтрифосфата
 Аденозин-3'-5'-циклофосфат 12
 5-Азатидин 219, 220
 Азот, метаболизм 153
 Активаторы транскрипции 81, 152,
 203—207
 Активация транскрипции 275, 295,
 296, 320
 Активность
 — АТРазная 49
 — белка автопротеазная 80
 — нуклеозидтрифосфатазная 156
 — полимеразная 48, 49, 50, 263
 — ферментативная 9, 219
 — хеликазная 271
 — 3'-экзонуклеазная 47—51, 77
 — 5'-экзонуклеазная 57, 65, 77, 272
 — эндонуклеазная 312
 Алкилирование оснований 74, 81
 Алкогольдегидрогеназа 197, 201—
 203, 224
 Альфа-вирусы 321—322
 Аминоацил-тРНК 169
 Аминоацил-тРНК-синтетаза 169
 α -Аманитин 136
 Амплификация 70, 200, 207, 226
 Антибиотики 113, 119, 122—125, 135
 Антигенные вариации 102
 Антигеном 322, 327
 Антигены 101, 219, 271, 299—302
 Антитерминатор 296
 — N-специфический 292, 294
 Антитерминаторная шпилька 160,
 161
 Антитерминация 161—162, 292, 295,
 297
 Апуринизация оснований 73, 74, 76
 Аргинин 217, 235—238
 Аренавирусы 328
 AP-сайт 73—74, 94
 Археобактерии 137
 AP-эндонуклеаза 76—77
 Атака нуклеофильная 46, 55
 АТРазы 52, 55, 89
 Аттенюаторы 158—161, 296
 Аутогенная регуляция 295, 301
 Аутокатализ 163, 166
 Аутосплайсинг см. Самосплайсинг
 Афидиколин 50
 Ацетилирование гистонов 254

 Бактериальные
 — ДНК-полимеразы 48—50, 54
 — плазмиды 110—112
 — репликоны 68—69
 — хромосомы 15, 54, 67, 104—109, 115
 Бактерии. См. также Транспозоны
 бактерий
 — генетическая изменчивость 126—
 129

- грамположительные 111
- репарация индуцируемая 78, 84, 94
- Бактериофаги 6, 15, 30, 78—84, 106, 107, 115, 136, 152
- рестрикция 129—133
- Белки 7, 8, 29, 47, 54, 111, 116, 117.
См. также Взаимодействия белок-белковые, белок-нуклеиновые
- активаторы 142, 144, 152, 196, 198, 230
- катаболитных оперонов (БАК) 148—150, 161
- активность автопротеазная 80
- антитерминаторы 162
- в репликации ДНК 52—53
- вирус-специфические 262—267, 290, 309, 318—327, 331
- гистоноподобные 56
- индукторы 304
- инициаторы 61, 62, 67
- истинно-ранние 297
- кинетохорные 72
- негистоновые 238, 242, 250, 257
- отсроченно-ранние 297
- ранние 304
- регуляторные 7, 9, 142, 144, 149, 152—154, 195, 211, 314
- — транскрипции 301
- — хроматина 249—251, 254, 257, 258
- репликации *E. coli* 56
- репрессоры 66, 80
- рибосомные 154, 165
- синтез 7, 8, 44, 80
- теплового шока *см.* Шок тепловой
- терминальные 264, 282, 320
- фагоспецифичные 265, 274, 275, 279, 287, 293, 296
- факторы транскрипции 195—203, 207, 209—212
- RecA 89—91, 94, 98
- RecBCD 90—91
- SSB 52—57, 82, 91, 92
- Белковая инженерия 9
- Белковый фактор тау 155
- Бирнавирussy 329

- Блоки
- геномные 189
- нуклеотидные в энхансерах 204
- Буньявирусы 327—328
- Валин 235
- Введение метки 19
- Векторы клонирования 63
- Взаимодействия. *См. также* Стэ-кинг-взаимодействия
- БАК с ДНК 148—149
- белок-белковые 53, 146, 234, 283
- белок-нуклеиновые 329
- внутримолекулярные 29
- — кооперативные 29, 152
- — типы 41—42
- гистона с фибриллой хроматина 255—256
- гистонов специфические 241—242
- ДНК с гистонами 235
- дуплексов 96
- комплементарные 37, 167, 178, 222
- координационные 46
- липких концов 281
- межнуклеотидные 22—26
- межплоскостные 26, 29, 37
- молекул репрессора 147
- рекомбинационные 96, 101, 102
- репрессора с оператором 144—148
- РНК I и РНК II 65—66
- РНК-полимеразы с промотором 154
- Вилки репликативные 49, 52, 53, 56—61, 65, 68, 91, 93, 116, 117, 234, 257, 263, 266, 271, 273, 275, 278
- схема организации 58
- Вирионы 261, 291, 299, 315, 316
- Вироиды 35
- репликация РНК 329—331
- Вирус
- аденоассоциированный 268, 270
- осповакцины 306—308
- полиомиелита 320, 331
- Вирусные
- геномы 20, 299, 308—316. *См. также* ДНК-геномы вирусов, РНК-геномы вирусов

- — интеграция с клеточными 283—289
- — репликация 260—332
- — транскрипция 260—316
- ДНК 15, 298—304, 312, 313
- РНК 19, 309, 314, 323—331
- Вирусoidы, репликация РНК 329—331
- Вирус-специфическая
 - ДНК 285, 287, 311—313
 - ДНК-полимераза 315, 316
 - РНК 274, 303, 306, 320
 - РНК-полимераза 299, 307, 329
- Вирус-специфический
 - РНК/ДНК дуплекс 311
 - фактор 297
- Вирусы 35, 42, 84, 110, 128—132, 204. *См. также* Аденовирусы, Альфа-вирусы, Вирус аденоассоциированный, Вирус полиомиелита, Гемини-вирусы, Паповавирусы, Парвовирусы, Поксвирусы, Эпштейна-Барр вирус
 - онкогенные 9
 - прокариот 290—299
 - репликация и транскрипция геномов 260—332
 - эукариот 299—308
- Внехромосомные ДНК 10
- Галактоза 230
- β -Галактозидаза 217
- Гамма-облучение 84
- Гель-электрофорез 7, 10, 16, 17, 19, 30, 40, 238, 240, 244
- Гемини-вирусы 260
- Генетическая изменчивость бактерий 126—129
- Геномная РНК *см.* РНК-геномы вирусов
- Геномы 101, 102, 108, 109, 129. *См. также* ДНК-геномы вирусов, РНК-геномы вирусов
 - аденовирусов 303—306
 - воспроизведение 47
 - ДНК-содержащих вирусов 289—308
 - млекопитающих 200, 233
 - обезьянного вируса 40, 299—303
 - нитчатых фагов 290—291
 - осповакцины 306—308
 - ретровирусов 308—315
 - — структура 308—309
 - РНК-содержащих вирусов 317—331
 - фага лямбда 274—276, 291—296
 - — — интеграция 283—285
 - — М-13 19
 - — Мю 285—289
 - — — схема интеграции 286
 - формирование элементов 223
 - число оснований 233
 - эукариот 185—195
 - — подвижные элементы 221—233
- Гены
 - активность 66, 107
 - высокомолекулярные 109
 - гомеостатические 215—218
 - иммуноглобулина 108, 109
 - инверсии 106—108
 - истинно-ранние 297, 298
 - кодирующие 54, 76, 81, 82, 84, 90, 101, 107, 111, 112, 131, 135
 - — белок-активатор 149, 151
 - мозаичные 9, 172, 173, 194
 - плазмид 113, 124
 - направляющие развитие организмов 212—218
 - поздние 303
 - ранние 299, 303, 304
 - репликационные 56—61, 64, 65, 273
 - рибосомные 207—209, 251—254, 256
 - секвенирование 20
 - строение экзон-интронное 9
 - структура первичная 7
 - теплового шока *см.* Шок тепловой
 - транскрибируемые РНК-полимеразой II 196—203
 - устойчивости к антибиотикам 113
 - фаговые 105—108, 293, 297
 - экспрессия *см.* Экспрессия генов
 - эукариот 185—221

- регуляторные элементы 196—203
- Гетеродуплекс 81, 88
 - асимметричный 100
- Гетеродуплексные участки 85—87, 99, 100
- Гетерохроматин 256
- Гетерохроматические участки хромосом 189
- Гибридизация
 - нуклеиновых кислот 186—187
 - РНК 65, 66
 - с клонированным фрагментом 243
- Гибриды ДНК и РНК 15, 62, 64, 139, 155, 173
- Гидролиз
 - аденозинтрифосфата 49—54, 78, 89, 91, 132
 - ДНК 15, 243
 - петель в хромосомах 247
 - РНК-затравки 57, 62
 - связей 13, 39, 45
 - хроматина 253
- Гипоксантин 73
- Гираза см. ДНК-гираза
- Гистидин 211, 237
- Гистоны 109, 234—238
 - аминокислотная последовательность 236
 - взаимодействие с фибриллой хроматина 255—256
 - — специфическое 241—242
 - генов транскрипция 253
 - карта линейного расположения 240, 242
 - структурная единица 237
 - типы 235
 - фосфорилирование 237, 238
- N-Гликозиды 11
- Гликозилирование ДНК 278
- O-Гликозилирование 193—194
- Гликопротеиды 101, 102, 109, 183
- Глицеральдегидтрифосфатдегидрогеназа 224
- Глобины 205, 206, 220, 224, 229
- Глутаминсинтетаза 153
- Глюкозо-6-фосфатдегидрогеназа 198
- Гомеоблок 217, 218
- Гомопиримидин 33
- Гомополинуклеотиды комплементарные 29
- Гонококки 101, 102
- Группы несовместимости 110, 111
- Гуаниловые РНКазы 14, 31
- Гуанин 5, 6, 11, 12, 21, 38, 73, 74, 143
- Гуанозин 12, 14, 166, 167, 177
- Гуанозинтетрафосфат 154
- Гуанозин-3'-фосфат 14, 174
- Двойная спираль
 - — ДНК 21—29, 31, 33, 34
 - — — антипараллельная 46
 - — — в нуклеосоме 239
 - — — левая 27, 32
 - — — плавление 29
 - — — полиморфизм 26—29
 - — — правая 27, 28, 32
 - — — расплетание 52—56, 61, 62, 138, 139, 265
 - — РНК 35—37, 41
- Двухжелевые ДНК и РНК 14, 35—37
- Деградация
 - ДНК 80—81, 128, 133
 - РНК в дуплексе 64
 - транскриптов 207
- Дезаминирование оснований 73—75, 83
- Дезоксиаденозин 12
- Дезоксигуанозин 12
- Дезоксигуанозин-3'-фосфат 12
- Дезоксинуклеозидтрифосфаты 18, 45, 46
- Дезоксипентоза 5
- 2-D-Дезоксирибоза 11, 12, 22, 73, 76
- Дезоксирибонуклеозид монофосфат 46
- Дезоксирибонуклеозиды 12
- Дезоксирибонуклеопротейд 263, 266
- Дезоксирибонуклеотиды 12, 58, 82, 264
- Дезоксицитидин 12
- Дезоксицитидин-3'-фосфат 12
- Деконденсация хроматина 251—253
- Денатурация ДНК 29—36, 143

- Денатурированная ДНК, кинетика реассоциации 186—188
- Депротенизация 10
- Дефекты ДНК, удаление 48
- Дефосфорилирование 153
- Дивергенция IS-элементов 124
- Дигидрокситимин 77
- Дигидротимин 77
- Дигидрофолатредуктаза 198
- Дидезоксинуклеозидтрифосфаты 18—19
- Диметилсульфат 16, 38
- α/α -Диплоиды 103
- Длина фрагмента, измерение 17
- ДНКаза 14, 140, 143, 253, 254
- ДНК-геномы вирусов 260—332
- двунитевые 264, 265
 - кольцевые однострунчатые 260, 262, 272
 - транскрипция 260—332
 - эукариот 299—308
- ДНК-гираза 34, 56, 60, 61, 92, 246, 272
- ДНК-N-гликозидаза 76, 77, 81
- ДНК-инсераза 76
- ДНК-лигаза 77, 206, 272
- в репликации ДНК 54—58
- ДНК-матрица 289
- ДНК-полимераза 6, 9, 20, 44—49, 63—79, 266, 272, 282. *См. также* Вирус-специфическая ДНК-полимераза
- активность 77
 - бактериальная 48—50
 - Кленова 48
 - процессивность 49—50
- ДНК-праймаза 50—57, 62, 71, 263—266, 271
- Додецилсульфат натрия 10
- Домены «пальцы» 211—212
- Дрожжевые
- геномы 69, 96, 207
 - клетки 96
 - РНК 41
 - хромосомы 15, 72, 96
- Дрожжи 167, 169
- гаплоидные 103
 - рекомбинация митотическая 95
 - штаммы гомоталлические 103
- Дрозофила, этапы развития 212—214
- Дуплексы 29, 44, 48, 62, 89, 94, 98, 111, 120, 269, 273, 284, 311. *См. также* Взаимодействия дуплексов
- гибридные 62—64
 - гомологичные 94, 95
 - расплетание 91, 266, 271
- Дупликация 225, 227
- генов 126, 127, 224
 - механизм образования 126
- Замены
- компенсаторные 35
 - нуклеотидные 35—37, 191
- Затравки 40, 44—47, 52, 63, 226, 263—266, 268, 271, 272, 279, 320, 326. *См. также* Праймер, РНК-затравка
- достраивание 51
 - множественные 263
 - нуклеотид-белковые 265—268
 - синтез ферментами 265
 - удлинение 50, 51, 79, 311
- Значащая нить ДНК 133, 134, 155
- Изомеризация полухиазм 87
- Изошизомеры 131
- Иммуноглобулины 108, 109, 176, 192, 202—205
- Инверсия 191
- концов цепи 269, 270
 - сайт-специфическая 107
 - сегментов ДНК 106—108, 114, 120, 121, 126
- Ингибиторы
- бактериальных РНК-полимераз *см.* Рифампицин, Стрептолидин
 - репликации 66
 - синтеза ДНК 46
- Индуктор 142
- Индукция
- профага 295
 - SOS-системы 78
- Индуцируемый энхансер 205—206

77, 78, 89, 98, 111, 130, 202, 203, 310—322, 326
 — — — — ДНК 262, 267—270
 — — — — зрелой мРНК 179—181
 — — — — кэпирование 309, 315, 326—329
 — — — полиаденилирование 302, 308, 309, 315
 — — — процессинг 169—175
 Конъюгация 91, 111, 120, 125, 128
 — участка ДНК 17, 69, 79
 Корепрессор 142
 Коронавирусы 317, 323
 Коррекция синтеза ДНК 47, 48
 Кор-фермент 135
 Кофакторы 54, 177
 Кроссинговер 95—102, 109
 — двойной 97
 — неравный 126
 Ксантин 73
 Кэп 177, 178
 Кэпирование 309, 315, 326—329
 Кэп-сайт 174, 175, 195—203, 224, 227
Кэрнса схема репликации 271—272, 275
 Лигаза 170—172, 275
 — фага Т4, 54
 — *E. coli* 54
 Лиганды-эффекторы 153
 Лигирование концов разрывов 116, 272
 Лизогенизация 293
 Линкер *см.* Межнуклеосомная ДНК
Максама — Гилберта метод 15—20
 Малые РНК 10
 — — ядерные 172—184, 212, 231
 — — — созревание 175
 — — — типы 182
 Матрицы
 — двухнитевые 265—266, 289
 — кольцевые 296
 — конкатемерные 296
 — линейные 278
 — моноцистронные 298
 — одонитевые 44—46, 262—264, 266, 267, 273

— полицистронные 290
 Матричная нить ДНК 133
 Матричная РНК (мРНК) 179—184, 201, 223, 224, 302, 317, 318, 323—329
 — вирус-специфическая 306
 — деградация 293
 — зрелая 173, 299, 305, 306
 — моноцистронная 163, 324, 326
 — поздняя 302—308
 — полицистронная 163, 164, 297, 298, 308, 331
 — ранняя 300—303, 307
 — субгеномная 318—322, 327, 332
 — типы 182
 Матураза 276
 Межнуклеосомная ДНК 243—245, 257
Мезелсона — Рэддинга модель 99—100
 Мембраны клеточные 68—69
 Металлотнионин 200, 201, 205, 206
 3-Метиладенин 75
 Метилаза 82, 119, 129—132, 219
 O⁶-Метилгуанин 74, 75
 7-Метилгуанин 12, 75
 Метилирование ДНК 82, 83, 119, 120, 123, 130—133, 254, 258
 — — как способ контроля 218—221
 Метилметансульфонат 76
 Метилтрансфераза 75, 81
 5-Метилцитозин 12, 218, 254
 Метки
 — радиоактивные 17—20
 — флуоресцентные 19
 Механизм
 — антигенных вариаций 102
 — вторичного разматывающегося рулона 274—277, 281
 — генетической нестабильности 120—122
 — генной конверсии 95
 — индукции SOS-системы 80, 81
 — инициации, самозатравочный 263—265, 268—270
 — интеграции плазмид 127
 — коррекции 47—48
 — мономеризации ДНК 105

- образования дупликаций 126
- перемещения мобильных элементов 115—120, 230
- разматывающегося рулона 272—274, 281
- регуляции репликации 63—68
- репликативной транспозиции 116
- транскрипции, молекулярный 289—308
- транспозиции 115—120
- упорядоченной сегрегации 68—69
- Миграция ветви 86, 87, 91
- Микрогетерогенность конформационная 29
- Микроорганизмы патогенные 101, 102, 107, 124, 128. *См. также*
- Вирусы
- Митохондрии 10, 15, 51, 136, 167, 186
- Михаэлиса — Ментен* уравнение 170
- Множественные тандемные копии *см.*
- Конкатемеры
- Мобильные генетические элементы 109—129
- механизм перемещения 115—120, 230
- транспозиция 222
- Модификация химическая. *См. также* Метилирование, Полиаденилирование
- ДНК 16, 17, 129—133, 143, 153
- конца транскрипта 181
- РНК 38—40, 163
- типы систем 129—133
- Молекулы ДНК
- двухнитевые 260, 261, 268, 269
- кольцевые 129, 260, 262, 271—274
- однонитевые 56, 65, 268
- Молекулярная гибридизация 30
- Молекулярное клонирование 191
- Мономеризация
- ДНК фага P1 105
- репликонов 105
- Мультигенные семейства 109, 190
- Мутагенез 80, 232
- горячие точки 83
- спонтанный 76
- Мутагены алкилирующие 81
- Мутанты *E. coli* 77, 79
- Мутации 54, 58, 62, 72, 73, 78, 80, 82, 92, 124, 132, 146, 154, 155, 199, 214, 215, 230, 232
- гомеотические (гомеиозисные) 215
- инсерционные 123, 126
- промоторные 140
- Направленность репликации 61, 63
- Нарушения ДНК 73—75, 94
- — прямая реактивность 75
- Нестабильность генетическая, механизм 120—122
- Ник-трансляция 48
- Нуклеазы 11—15, 33, 34, 39, 40, 91, 133, 169, 227, 239, 249, 253, 256, 257, 266
- из яда змей *см.* Яд змеиный
- микрококковые 243—244, 253
- Нуклеиновые кислоты 8, 10, 22—26
- — обмен 9
- — очистка от белков 10
- — структура 6, 10—42
- — типы 5
- Нуклеозидтрифосфаты 19, 138, 160, 320
- активность 156
- Нуклеозид-5'-фосфаты 14
- Нуклеозиды 6, 11, 12, 25, 319—320
- анализ конформационный 23—26
- минорные 12
- пиримидиновые 23
- пуриновые 23
- Нуклеомеры 245
- Нуклеоплазмин 255, 258
- Нуклеопротейды 10, 326
- Нуклеосомная ДНК 245
- Нуклеосомы 9, 234—242, 256, 257
- минимальные 244
- организация фибрилл 242—244
- пространственная модель 239
- реконструкция 241
- транскрипция 254—256
- фейзинг 242
- Нуклеотиды, номенклатура 11, 12

- Облучение ультрафиолетовое 73—80
- Обратная транскрипция *см.* Транскрипция обратная
- Однотяжевые
 - ДНК 33, 52, 55
 - РНК 35, 37
 - — структура 37—42
- Оказки фрагменты 53—58
 - — сшивание 56
- 5-Оксиметилцитозин 133, 262, 278
- Октамер гистоновый 241—242, 258
- Олигонуклеотиды 14, 20, 27, 48, 80, 138, 320
 - синтез 40
 - фосфорилирование 14
- Олигорибонуклеотид-затравка 263
- Онкобелки 9
- Онкогены 9
- Операторные зоны 295
- Операторы транскрипции 142, 144
- Опероны 135, 160—165
 - катаболитные 148—151
 - — репрессоры 150—152
 - триптофановый 158—161
- Органеллы клеточные *см.* Митохондрии, Хлоропласты
- Ориджин
 - переноса 111
 - расплетание 61, 62
 - репликации 60—70, 12, 120
 - специфичность 61
- Основания
 - алкилирование 74, 81
 - апуринизация 73, 74, 76
 - гетероциклические 11, 24, 26
 - дезаминирование 73—75
 - интеркалляция 41, 42
 - минорные 12
 - модифицированные 262
 - пиримидиновые 22
 - — расщепление 16
 - пуриновые 22, 41
 - — метилирование 16, 131
- Остатки нуклеотидные
 - — идентификация 40
 - — конформация 22—26
 - — минорные 11
- — модификация 38—39
- — терминаторные 18
- — фосфатные 24
- Очистка ДНК и РНК 10
- Палиндромная шпилька 181
- Палиндромы 33
- Паповавирусы 260, 271—272
- Парвовирусы 262—265, 269, 270
- Пары
 - комплементарные 24—26, 37
 - третичные 41
 - уотсон—криковские 24—26, 37, 41
- Паузы транскрипции 140, 155, 157, 160
- Пентоза 5, 11
- Перекрытие оснований 26
- Перестановки кольцевые 261, 280
- Петли
 - ДНК 37, 66, 147—153, 173, 181
 - лидерные 158—161
 - суперспирализованные 246
 - энхансеров 204
- D-Петля 65, 89, 90, 92
- Пикорнавирусы 319
- Пилин 101, 102
- Пиримидины 5, 11, 12, 16, 76, 242, 257, 262. *См. также* Тимин, Цитозин
 - димеры 75, 80
- Пирофосфатаза 46
- Плавление двойной спирали ДНК
 - 29—31, 143
 - — — кооперативное 29, 138
- Плазмиды 61, 63—69, 96, 104, 106, 109—115, 120
 - бактериальные 15, 110—112, 122—129
 - интеграция 127
 - колициногенные 125
 - конъюгативные 125, 128
 - мобилизуемые 111
 - репликация 110
 - трансмиссивные 111
- Повторы 301
 - длинные 225—230
 - инвертированные 112—115, 122—124, 231—233, 292, 294, 312

- — концевые 261, 262, 267
- короткие 225—227
- нуклеосомные 243
- прямые 285, 313
- — концевые 261, 277, 278, 280, 309—311
- тандемные 189, 209, 235
- Позиционная информация 214
- Поксвирусы 262
- Полиаденилирование 165, 179—181, 197, 223—230, 237, 302, 305, 308, 309, 315
- Полимеризация 46—48
- Полиморфизм двойной спирали ДНК 26—29
- Полинуклеотидкиназа 17
- Полинуклеотиды 7, 14, 19, 25, 27, 34, 47, 48
- Полипептиды
 - активность 48
 - вирус-специфические 321
 - в ДНК-полимеразе 48, 57
 - протеолиз 48
 - процессинг 8
- Полирепликационная организация 69
- Полирибонуклеотиды 163
 - гидролиз 15
 - природные 40
- Полуметилированная ДНК 82—83, 120, 123, 258
- Полухиазма 85—89. *См. также* Изомеризация полухиазм
 - образование 89—92, 99
 - разрезание 86—89, 92, 94, 96—100
- Порядок зацепления цепей ДНК 31
- Последовательности
 - гомополимерные 322
 - — дублицированные 112, 122
 - — вирус-специфические 325
 - — гомологичные 103, 109
 - — гомопиримидиновые 33
 - — инвертированные 33
 - — промоторов прокариот 140—142
 - — — канонические 141
 - — регуляторных элементов генов 196—200
- — РНК, определение 19—20
- — самокомплементарные 268
- — усилителей транскрипции 204
- Праймаза 274, 275, 282. *См. также* ДНК-праймаза
- Праймер 19, 20. *См. также* Затравка
- Праймосома 54—58, 62—65, 263
- Предшественники тРНК, процессинг 168—172
- Прибнова блок 141
- Провирус 284, 289, 313—315
- Прокариоты 10, 35, 110, 135—163, 231, 257
 - вирусы 290, 317
 - доменная организация хроматина 246
 - промоторы 142—154
 - процессинг 163—165
 - — предшественников тРНК 168—172
 - регуляция транскрипции 249
- Промоторы 7, 133—154, 250, 254, 290, 293. *См. также* Транскрипция с промоторов
 - активация 296
 - галактозного оперона 150
 - истинно-ранние 296
 - катаболитных оперонов 148—150
 - отсроченно-ранние 296
 - поздние 296, 297
 - прокариот 142—154
 - — активация репрессором фага 146
 - — активность негативная 142
 - — — позитивная 142
 - — — регуляция 142—154
 - ранние 297
 - структура 140
 - эукариот 196—197
- Проретровирусы 224, 227
- Протаминны 238
- Протеаза 144, 153
- Протеолиз 48, 318, 331
- Профаг 81, 106, 107, 115, 284, 285, 291, 294, 295
- индукция 295
- Процессивность синтеза ДНК 49—50

- Процессинг 207, 210, 224, 226, 305
 — конца транскрипта 178—181
 — определение 163
 — первичных транскриптов 163
 — предшественников тРНК 168—172
 — у прокариот 163—165, 167
 — у эукариот 165—168
 Псевдогены, образование 223—225
 Псевдоузлы 37
 — схема образования 42
 Псевдоуридин 12
 Пурины 5, 11, 12, 16, 41, 141, 242, 257. *См. также* Аденин, Гуанин
- Разделение цепей 17
 Размыкание пуринового кольца 74—75
 Разрыв
 — двухцепочечный 79, 94, 95, 97, 103, 116—120, 130, 289
 — иницирующий 97, 266
 — одноцепочечный 99, 111, 116
 Расхождение хромосом 108
 — межнуклеотидных связей 13, 39, 40
 — полинуклеотидных цепей 15, 20
 Раунд репликации 60, 61, 64, 66—72
 — инициации 56, 66, 67, 68, 274
 Реактивация прямая 75, 76
 — Уэйгла 78, 79
 Реассоциация цепей ДНК 31
 — у эукариот 186—188
 Реверсия 121
 Ревертаза *см. Транскриптаза обратная*
 Регуляция
 — синтеза белков 7
 — транскрипции 209—212, 290
 — — ДНК-геномов вирусов прокариот 290—299
 — — фага лямбда 294, 301
 — — — аутогенная 295
 — экспрессии генов 181, 198, 212, 293, 302, 303
 — — — металлами 197
- — — путем альтернативного сплайсинга 182—184
 — — генома 299
 Резолваза 117, 118, 123
 Рекомбинантные ДНК 9, 34—35, 215
 Рекомбинационная инициация, схема 279
 Рекомбинация 79, 94, 266
 — геномов коронавирусов 323
 — гомологичная 81, 84—89, 93, 98—101, 103—105, 108, 112, 122, 126, 127, 229
 — — специализированные системы 102—104
 — иницируемая двухцепочечным разрывом 96
 — мейотическая 98—100, 108
 — митотическая 95—99
 — сайт-специфическая 104—108, 117, 118, 283—285
 — — механизм 285
 — частота 128
 — эволюционная роль 108—109
 Релаксированное состояние ДНК 31—34
 Ренатурация ДНК 29—36, 52, 91
 — — скорость 30—31
 Рентгеноструктурный анализ 7, 21, 23, 26, 29, 38, 40, 239, 240
 Репарация ДНК 48—51, 73—84, 95, 202, 204
 — двухцепочечных разрывов 79, 94, 95, 97, 103, 116—120, 130
 — индуцируемая 78—81
 — неспаренных нуклеотидов 81
 — пострепликативная 79, 93, 94, 108
 — рекомбинационная 78, 94, 95, 101, 120
 — темновая 75
 — эксцизионная 76—78
 — SOS 78—79
 Репликаза 319, 320
 Репликативная
 — вилка *см. Вилка репликативная*
 — транспозиция 285—289
 — форма ДНК 273, 282
 Репликация 8, 18, 34—35, 44—72,

- 81, 82, 89, 119, 120, 219, 234; 266.
См. также Вилка репликативная,
 Раунд репликации
- геномов вирусов 260—332
 - — — гриппа 325—327
 - — РНК-содержащих вирусов 317—331
 - — фага лямбда 274—276, 277—280
 - — — Мю 285—289
 - — — Т-серии 277—280
 - ДНК-геномов 260—289
 - — с однокитовой кольцевой ДНК 272—274
 - двунаправленная 61
 - дихотомическая 67—68
 - избирательность 56
 - кольцевых ДНК 58—59
 - мобильного элемента 114, 116, 117
 - однонаправленная 61, 63
 - основные принципы 51—54
 - — схемы 266—280
 - плазмид 110
 - полуконсервативная 44, 45, 263, 268, 278, 287
 - расплетания двойной спирали ДНК 52—53
 - регуляция 63—68
 - РНК вирионов и вирусондов 329—331
 - РНК-геномов двухкитовых 328—329
 - с использованием конкатемеров 277—280
 - скорость 47, 57
 - схема вторичного разматывающегося рулона 274—276, 281
 - — Кэрнса 271—272, 275, 281
 - — разматывающегося рулона 272—274, 281
 - топологические проблемы 56, 59—60
 - хроматина 257—258
 - хромосомы *E. coli* 61—63
 - эукариот 69—72
- Репликоны 60, 63, 64, 67, 70, 110, 112, 116—124, 127
- кольцевые 61, 105, 106
 - мономеризация 105
 - сегрегация 68—69, 105
 - слияние 114
- Реплисома 58, 61
- Репрессия 67, 144, 150, 293
- Репрессоры 142, 144, 293, 295
- катаболитных оперонов 150—152
 - концентрация 147—148
 - транспозиции 123
 - фагов 81, 144—148
 - — лямбда 148—152
- Рестриктаза 13—16, 129—133, 225, 248
- Рестрикция 79
- бактериофагов 129—130
 - ДНК 129—133
 - системы, типы 130—132
- Ретровирусы 308—315
- Ретрогены 224
- Ретропозиции ДНК-копий 225
- Ретропозоны 222—227
- Ретротранспозоны 227—230
- Д-Рибоза 11, 12, 22, 42, 73
- Рибозимы 170
- Рибонуклеозидтрифосфаты 133
- Рибонуклеозиды 12, 15, 20, 36—38
- Рибонуклеопротеиды 178, 181, 238, 323
- Рибонуклеотиды 51, 163
- Рибосомные РНК 7, 19, 190, 209, 210, 251
- — антитерминация в генах 161—162
- Рибосомы 7, 8, 38—42, 154—159, 207, 237
- Рифампицин 135, 136
- РНК
- информационные 7, 19
 - расщепление цепи 15
 - структура макромолекулярная 35—42, 63, 66
- РНКазы 14, 15, 38, 39, 56, 62—65, 157, 164, 165, 169, 172, 293, 298, 308, 311
- РНК-геномы вирусов 317—331
- — виды 317

- — двухнитевые 328—329
- — одонитевые 317—323
- — репликация 317—331
- — транскрипция 317—331
- РНК-затравка 51—57, 62—65, 70, 267, 272, 310
- гидролиз 57
- первичная 265
- РНК-лигаза 20, 331
- РНК-полимераза 7, 46, 62, 64, 133—160, 195, 197, 202, 234, 251—256, 292, 293, 296—299, 307, 323, 324, 327
- вирус-специфическая 299, 329
- ДНК-зависимая 263—266, 277—279, 290, 291, 307, 330
- модификации 297
- поздняя 297
- субъединицы 297
- РНК-зависимая 319—323
- фагоспецифическая 277, 298
- I 238, 251, 252
- II 238, 250, 253, 254, 301, 303, 307, 314
- III 226, 238, 303

Салмонелла, смена фаз 107—108

Самокомплементарность 27, 29, 262

Самосплайсинг 166, 167, 177, 331

Сателлитные ДНК 188—190

Сахара 12, 29, 75, 148

Сверхспирализация 31—35, 61, 62, 204, 271, 272

- отрицательная 32—34, 62, 92, 139

Сворачивание полипептида 8

Связи

- водородные 25, 41, 42, 86
- — гликозидные 11, 75, 76
- межнуклеотидные 11—15, 40, 166
- — гидролиз 13
- фосфодиэфирные 11, 39, 46, 167, 174, 320
- — образование 55, 177, 264
- — расщепление 40, 77, 116, 131

Сегментация генома 332

Сегменты инвертируемые 106—108

Сергегация

- плазмид 124
- постмейотическая 85, 86, 88, 99, 100
- репликонов 105
- упорядоченная 68, 69, 72, 99, 104, 106

Секвенирование 7, 15, 16, 18, 20, 40

- по Максаму — Гилберту 15—19
- по Сэнгеру 17—20

Сигма-субъединицы 135—142, 161

- минорные 135, 153
- РНК-полимеразы 152—154

Синапс 98

Синаптонемный комплекс 98, 99

Синтез

- ДНК матричный 45—48
- первичных транскриптов 303—305
- — прерывистый 53—54
- — процессивность 49
- — репарационный 269
- — точность и коррекция 47—52

Содержание ДНК в разных организмах 185

Сортировка экзонов 193—195

Спаривание

- гомологов 98
- липких концов 280
- нуклеотидных остатков, комплементарное 38, 65, 66, 85, 93, 133, 155
- — — нарушения 81
- у дрожжей, типы 103, 104
- хромосом гомологическое 98

Спейсеры 207—209, 251, 252

α -Спираль 144, 145, 217, 235

Сплайсинг 9, 166—181, 302, 308, 315, 326

- альтернативный 182—184, 191, 308, 217
- нарушения 175
- стадии 177
- экзонов 172—174, 197

Сплайсосома 178, 179

Степень репрессии 147, 152

Стрептолидигин 135, 136

Стэкинг-взаимодействия 24, 26, 29,
37, 41, 42, 86
Субвирусные частицы 329
Суперглобулы 245
Сшивки ковалентные
— — белков с ДНК 250, 264
— — гистонов с ДНК 254
— — межнитевые 83, 240
Сэнгера метод 17—20
SOS-системы 78—80

Теломеры хромосом 71, 72, 101
Терминаза 276
Терминаторная шпилька 155—161
Терминаторы 7, 133
— прокариот 154—182
— регулируемые *см.* Аттенюаторы
Терминация
— в отсутствие белковых факторов 155
— транскрипции 210, 251
— трансляции 197
— цепи РНК 134, 137, 158, 179—181
— р-зависимая 157, 161
Тимидилатсинтаза 167, 194
Тимидин 12
Тимидин-5-фосфат 12
Тимин 5, 6, 11, 12, 21, 74, 75, 262
— димер 77, 79, 94
Т-клетки 108
Транскрипт 201, 225
— процессинг 3'-конца 178—181
Транскриптон 133—135
Транскриптаза обратная 20, 40,
222—227
Транскрипты первичные
— — процессинг 302—303, 305—
306
— — синтез 300—305
Транскрипция 7, 8, 61, 62, 66, 123,
133—162
— активация 81, 230, 275
— геномов вирусов 260—332. *См.*
также Геномы вирусов
— обратная 40, 195, 222—227, 308—
312

— регуляция 154—162
рибосомных генов 207—209, 252
— термины 133
— цикл 134, 137—142
— — стадии 175
— эукариот 195—221
Транслокации 114, 126
Трансляция 7, 8
Транспозаза 112—120, 123, 124, 231,
232
Транспозиция 112—115, 118
— молекулярный механизм 116—
119
— репликативная 118, 120—122
— консервативная 119
— Р-элемента 231—232
Транспозоны
— бактерий 112—127
— определение 113
— перемещение, механизм 118—119
Транспорт
— молекул ДНК в клетку 128
— РНК 8, 170
Транспортные РНК (тРНК) 190,
209, 210, 310
— — процессинг предшественников
169—172
Транс-сплайсинг 184, 307
Трансфераза теломерная 71
Трансэтерификация 166, 167, 176,
177, 195
Триметилгуанидин 177
Триозофосфатизомераза 176, 192
Трипаносомы 101, 109
Триплеты 41
Трипсин 241, 245
Тэта-структуры 61

Узелки рекомбинационные 98
Уотсона — Крика модель 21—24, 44.
См. также Пары уотсон-криков-
ские
Урацил 5, 12, 38, 73—75
Урацил-ДНК-гликозилаза 76, 83
Уридин 12, 23, 177
Уридин-5'-фосфат 10
Усилители

- транскрипции 196, 203—207
- — механизм действия 204—205
- Устойчивость к антибиотикам 113, 119, 122, 124, 125
- Уэйгла реактивация 78, 79
- Фаги 133, 261—263, 268, 272. *См. также* Бактериофаги, Профаг
 - лямбда 81, 104, 130, 157, 217, 261, 265, 274—276
 - — антитерминация 161—162
 - — интеграция генома 283—285
 - — репрессия 144—149
 - — репликация 277—280
 - — транскрипция генома 291—296
 - M13 19
 - Мю 106—108, 115, 117, 285—289
 - — интеграция генома 286
 - — нитчатые, транскрипция генома 290—291
 - Т-серии 92, 98
 - — репликация геномов 277—280, 282
 - — транскрипция генома 296—299
 - содержащие РНК 318—320
 - N4, транскрипция генома 299
 - фХ174 273—275, 282
- Фактор терминации 156, 157, 162
- F-Фактор 125, 128
- Фейзинг нуклеосом 242
- Фенилаланингидроксилаза 175
- Ферменты 6, 7, 13, 20, 117
 - активность экзонуклеазная 48
 - расщепление межнуклеотидных связей 13, 40
 - — полинуклеотидных цепей 15
 - репарации 83, 94
 - фотоактивирующие 75
- Фибриллы
 - нуклеосом, организация 242—244
 - хроматина 234, 235, 238, 242, 244
 - — взаимодействие с гистонами 255—256
 - — модели структуры 245, 255—256
- — разрушение 255
- Фибронектины 183
- Филогенетический подход 37—40
- Флагеллины 107, 108
- Форма ДНК 22, 26—29, 32—34, 98, 220, 257
- Формамидопиримидин 74—76
- Фосфатизомераза 224
- Фосфоглицераткиназа 198
- Фосфодиэстераза 14, 171
- Фосфорилирование 153, 177
 - в гистонах 237, 238, 248
- Фотолиаза 75
- Фотореактивация 75
- Фракционирование ДНК и РНК 10
- Хелаты 46
- Хеликаза 52, 54—61, 77, 78, 82, 91, 111, 265, 266, 275, 282
 - вирус-специфическая 271
- Хиазмы 99
- Хи-сайт 91
- Хлоропласты растений 10, 15, 51, 136, 167
- Холидея структура 85—88, 92, 96, 100 *См. также* Полухиазмы
- Хроматиды гомологичные 99
- Хроматин. *См. также* Фибрилла хроматина
 - активный, структура 251—254
 - гиперчувствительные места 256—258
 - деконденсация 251—253
 - доменная организация 246—247
 - интерфазный 252, 258
 - конденсация 244—245, 256
 - расщепление 243—244
 - репликация 257—258
 - структура 233—259
 - уровни компактизации 249
- Хромосомы 72, 98, 99, 220. *См. также* Теломеры хромосом
 - бактериальной клетки 10, 15, 67, 104—109, 115
 - — — механизм репликации 54—58
 - метафазные 247—249, 258
 - политенные 185, 251

- размеры 15
- расхождение 108
- спаривание гомологичное 98
- эукариот 69—72
- *E. coli* 61—63

Центромера 72, 99, 103

Цепи 20, 37. *См. также* Участки цепей

- двойной спирали ДНК 55—56, 78
- матричные 54—58, 64, 82
- полинуклеотидные 13, 15, 16, 23, 37, 45, 59, 71, 78, 93, 262, 265, 267 269

— — полярность 11

— — размеры 19

2'-3'-Циклофосфат 170—171

Цитидин 12

Цитидин-3'-5'-дифосфат 20

Цитозин 5, 6, 11, 12, 21, 38, 73—75, 129, 133, 218—220, 262

— дезаминирование 75

— метилированный 83

Цитоплазма 10, 163, 170, 179, 186, 193, 214, 219, 237, 241, 250, 251, 258

Чаргаффа правило 21

Шок тепловой 153, 199, 200, 205, 206, 250, 252, 255

Шпильки

— ДНК 264, 266, 268

— РНК 222, 226

«Эгоистичная» ДНК 122—126

Эксонуклеаза 14, 77, 90, 96, 104, 130, 164, 165, 169, 170, 180

Экзоны

— размеры 176

— сортировка 193—195

— экспрессия 182, 183

Экспрессия генов 9, 101—103, 107, 154, 163, 167, 178, 191, 194, 195, 200, 203, 210

— иммуноглобинов 205

IS-Элементы бактерий 112—127

Электрофорез 16, 143, 177, 243.

См. также Гель-электрофорез

Элонгация цепей ДНК 135, 319, 140, 266, 272

Эндоконформации 22—23, 26, 35

Эндонуклеазы 76, 78, 97, 165, 170

— АТФ-зависимые 77

— неспецифические 14

— рестрикции *см.* Рестриктаза

— сайт-специфические 103, 104, 133

Энзимология рекомбинаций у бактерий 88—93

Энхансеры *см.* Усилители экспрессии генов

Эписомы 127, 128

Эпштейна — Барр вирус 261

Эукариоты 10, 35, 50, 51, 54, 59, 60, 78, 81, 84, 95, 97, 101, 108, 110, 128, 133, 135—137, 163, 165, 184

— геномы 185—195, 221—233

— особенности репликации 69—72

— прерванные гены 191—195

— процессинг 165—168, 172—179, 180, 181

— — предшественников тРНК 169—172

— реассоциация ДНК 186—188

— регуляция транскрипции 195—221

— структура хромосом 233—259

Эффект положения 256

Эффектор 50, 132, 142, 154

Яд змеиный 14, 39, 200

Ядерный магнитный резонанс 38

Ядра клеток, исследования 234—235

ОГЛАВЛЕНИЕ

Предисловие	3
Введение	5
<i>Часть А. ВВОДНАЯ</i>	10
<i>Глава I. Структура нуклеиновых кислот</i>	10
1. Первичная структура нуклеиновых кислот	10
2. Макромолекулярная структура нуклеиновых кислот	20
Литература	43
<i>Часть Б. Биосинтез нуклеиновых кислот</i>	44
<i>Глава II. Репликация ДНК</i>	44
1. Матричный синтез	44
2. Основные принципы репликации	51
3. Репликация ДНК <i>E. coli</i>	54
4. Топологические проблемы репликации	59
5. Инициация раунда репликации	60
6. Сегрегация бактериальных репликонов по дочерним клеткам при делении	68
7. Особенности репликации ДНК эукариот	69
Литература	73
<i>Глава III. Репарация ДНК</i>	73
1. Нарушения, возникающие в ДНК	73
2. Прямая реактивация повреждений ДНК	75
3. Эксцизионная репарация	76
4. Индуцируемая репарация	78
5. Репарация неспаренных нуклеотидов	81
6. Метилирование ДНК и «горячие точки мутагенеза»	83
Литература	84

Глава IV. Рекомбинация	84
1. Гомологичная рекомбинация	84
2. Специализированные системы гомологичной рекомбинации	101
3. Сайт-специфическая рекомбинация	104
4. Эволюционная роль рекомбинации	108
Литература	109
Глава V. Плазмиды и мобильные генетические элементы бактерий	110
1. Бактериальные плазмиды	110
2. IS-элементы и транспозоны бактерий	112
3. Плазмиды и мобильные элементы бактерий и концепция «эгоистичной» ДНК	122
4. Плазмиды, мобильные элементы и генетическая изменчивость бактерий	126
Литература	129
Глава VI. Рестрикция и модификация ДНК	129
1. Рестрикция бактериофагов	129
2. Системы рестрикции	130
3. Борьба бактериофагов с системами рестрикции хозяина	132
Литература	133
Глава VII. Транскрипция	133
1. РНК-полимеразы	135
2. Цикл транскрипций	137
3. Регуляция активности промоторов прокариот	142
4. Регуляция транскрипции в терминаторах прокариот	154
Литература	162
Глава VIII. Процессинг первичных транскриптов	163
1. Процессинг у прокариот	163
2. Процессинг рРНК у эукариот	165
3. Процессинг предшественников тРНК у про- и эукариот	169
4. Процессинг РНК, синтезируемой с помощью РНК-полимеразы II, у эукариот	172
5. Регуляция экспрессии генов путем альтернативного сплайсинга	182
6. Транс-сплайсинг	184
Литература	184
Глава IX. Геном эукариот	185
1. Кинетика реассоциации денатурированной ДНК	186
2. Сателлитные ДНК	188
3. Умеренно повторяющиеся последовательности	190
4. Уникальные последовательности генома	190
5. Прерванные гены эукариот (экзоны и интроны)	191
Литература	195
	351

Глава X. Регуляторные элементы генов и регуляция транскрипции у эукариот	195
1. Регуляторные элементы генов, транскрибируемых РНК-полимеразой II	196
2. Энхансеры транскрипции генов, транскрибируемых РНК-полимеразой II	203
3. Транскрипция рибосомных генов с помощью РНК-полимеразы I	207
4. Регуляция транскрипции, осуществляемой РНК-полимеразой III	209
5. Гены, направляющие развитие организмов	212
6. Метилирование как способ контроля активности генов эукариот	218
Литература	221
Глава XI. Подвижные генетические элементы генома эукариот	221
1. Элементы генома, представляющие собой продукты обратной транскрипции клеточных РНК (ретропозоны)	222
2. Подвижные элементы с длинными концевыми повторами (ретротранспозоны)	227
3. Мобильные элементы, ограниченные инвертированными повторами	231
Литература	233
Глава XII. Структура хроматина	233
1. Гистоны	235
2. Негистоновые белки	238
3. Нуклеосома	238
4. Организация нуклеосомных фибрилл	242
5. Конденсация хроматина	244
6. Доменная организация хроматина	246
7. Метафазные хромосомы	248
8. Регуляторные белки хроматина	249
9. Структура активного хроматина	251
10. Судьба нуклеосом при транскрипции	254
11. Гиперчувствительные места хроматина	256
12. Репликация хроматина	257
Литература	258
Часть В. БИОСИНТЕЗ ВИРУСНЫХ НУКЛЕИНОВЫХ КИСЛОТ	
Глава XIII. Репликация и транскрипция вирусных геномов	260
1. Репликация генома ДНК-содержащих вирусов (ДНК→ДНК)	260
2. Транскрипция генома ДНК-содержащих вирусов (ДНК→РНК)	289
3. Репликация/транскрипция вирусных геномов, включающая синтез ДНК на матрице РНК	308
4. Репликация/транскрипция геномов РНК-содержащих вирусов (РНК→РНК)	317
Литература	332
Предметный указатель	334